A TWO-PART PERSONALITY MEASURE FOR USE AS A RESEARCH CRITERION

BY ALASTAIR HERON*

Medical Research Council Unit for Research in Occupational Adaptation

The present paper describes the development and evaluation of a 20 min. two-part personality measure suitable for use as a research criterion in a wide variety of situations. One part provides a score of emotional maladjustment based on twenty of seventy-four items, having an internal consistency of 0·81; the other part a score of sociability based on twelve of thirty-six items, having an internal consistency of 0·74. Evidence is presented for the validity of both scores, and for their independence of age and intelligence as assessed in adequately heterogeneous samples.

I. INTRODUCTION

Psychological research is not always directed to the evaluation of individual differences in intelligence or temperament, but experienced workers constantly encounter situations where such evaluation is essential as a control in the selection of experimental subjects. It is therefore a customary precaution to obtain a measure of intelligence in case it may prove to be related to the variables under investigation, but all too often a similar precaution is omitted where emotional factors might be a further source of variance. Such an omission may often be due to the lack of a suitably brief, reliable and valid measure. The present paper describes the development of such a measure, which takes the form of a two-part inventory which has during a period of over four years been used widely under various conditions in industry and elsewhere. Publication has been withheld until the information concerning its characteristics was considered adequate to justify making it available for use as a criterion measure suited to the needs of research workers with a wide range of interests.

Most readers will be familiar with many of the questionnaires and inventories which have been published during the past forty years; by now it is hardly possible to produce a new item. Such inventories have been variously named, but in one way or another the majority can be described as concerned with emotional maladjustment, some aspect of 'sociability' or 'introversion-extraversion' coming a close second.

The first part of the present inventory is intended as a measure of emotional maladjustment; a determined effort has been made to provide in the second part a measure of sociability which is as far as possible independent of the first measure. As historically the development of the first part preceded that of the second, and provided evidence which is still relevant to an evaluation of the final product, a brief summary of this development stage is necessary at the outset.

II. DEVELOPMENT OF FIRST PART

The original draft consisted of seventy-four items, and this has been maintained as the total throughout the whole programme of development. This was made up of three types of item: thirty intended as indicators of emotional maladjustment when marked in the significant direction; thirty concerned with physical health, intended principally as

* Now Deputy-Director, Medical Research Council Group for Research on Occupational Aspects of Ageing, University of Liverpool.

buffers for the emotional items; and fourteen drawn from the so-called 'Lie Scale' of the Minnesota Multiphasic Personality Inventory (MMPI). The thirty-six items in the first group were drawn mainly from the Maudsley Medical Questionnaire (Eysenck, 1952) and from Winne's study (1951) of selected items from the MMPI, together with a few items from Cattell's factor analysis studies and from U.S. Army questionnaires. All items found in the form of questions were re-cast as statements, some were altered from positive to negative form or vice versa, and quite a few had to be modified slightly because of unfamiliar American idioms.

In this original form the inventory was used in a large-scale study of unskilled male and female workers in electrical engineering factories. The samples involved consisted of men ($N=100$) and women (N's$=100$ and 124) between the ages of 18 and 60. No significant difference was found for emotional maladjustment scores between male and female samples, and all correlations with age and intelligence were statistically insignificant. Although every one of the original thirty items had been shown by the source authors to distinguish significantly between groups of 'neurotic' and of 'normal' respondents, it seemed desirable to obtain evidence of internal and external validity for the total score when used in the existing circumstances and in company with the buffer items referred to. In one sample of 124 women who had been subjects in a study involving a large number of personality tests (Heron, 1955*b*) the thirty-item total score obtained the highest loading (0·64 after rotation) on a factor readily identifiable as emotional instability or maladjustment. In this study the split-half (odd-even) self-correlation, corrected to full length, was 0·82. When the distribution of scores for fifty female neurotic in-patients was compared with that for a group of twenty-one female unskilled workers passed by a psychiatrist colleague as 'apparently free from signs of neurotic illness', only 12% overlap was found. It therefore seemed worth while continuing to use the full seventy-four-item inventory, while contemplating the possibility of re-keying the score for emotional maladjustment in the light of fresh data and further analysis. The only alteration of actual items was therefore carried out at this stage, six items being dropped (four from the 'Lie Scale' and two from the 'physical health' group), all of which were replaced by six fresh items likely to measure emotional maladjustment, thus raising the total score for this to thirty-six.

In this form, not subsequently altered, the inventory was employed in a fresh study of 214 new entrants as male omnibus conductors (Heron, 1955*a*; Venables, 1953). The split-half self-correlation, corrected to full length, proved to be 0·83 whether one used odd *v.* even items or first half *v.* second half as the basis of calculation. Correlations with age and intelligence remained insignificant, and the loading in a large heterogeneous battery on a factor of emotional maladjustment was again the highest at 0·53. Overlap of distributions between seventy-five symptom-free men and fifty neurotic in-patients was 16%. It was possible to obtain re-test data for seventy-eight of the original 214 men after an interval of 6 months. As these seventy-eight were of course among those who had not resigned or been discharged from the job, the re-test correlation was likely to be shrunk due to restriction of range if emotional maladjustment scores happened to be related to leaving behaviour. This proved to be the case, and accordingly the obtained coefficient was corrected by means of the two variances concerned, giving a final estimate of 0·75. It will be noticed that this does not differ greatly from the split-half correlation of 0·83, suggesting that 'function fluctuation' is relatively slight.

Before describing final development work on the first part of the inventory, it will be necessary to describe the development of the second part which took place at this stage.

III. DEVELOPMENT OF SECOND PART*

The item-source for a measure of sociability used on this occasion was Guilford's '*R*' (rhathymia) scale. On the basis of material obtained by Hildebrand (1953) the seventy-one items were analysed for frequency of response on the part of two contrasting groups of hospital patients, twenty-five diagnosed as hysteric, twenty-five as anxiety states. After elimination of items too American in idiom for general use, thirty-six were selected which best discriminated between the two groups. It will be observed that an assumption was made here in order to overcome the apparently insoluble problem of obtaining criterion groups of extreme 'introverts' and extreme 'extraverts' in the general population. The assumption is that a man diagnosed as an 'hysteric' is more likely to be a neurotic 'extravert' than a neurotic 'introvert', and contrariwise for anxiety states. This assumption finds support in the work summarized by Eysenck (1952, 1953). The term 'sociability' has been selected for this part of the present inventory mainly because it is somewhat less likely than is 'introversion-extraversion' to be narrowly identified with any particular school of thought in contemporary psychology.

These thirty-six items were then used as a second part of the inventory, administered after an interval during which some other test or activity was inserted. The need for such a break was found to be generally desirable from observations made during the administration of the seventy-four-item first part to several hundred industrial respondents of both sexes. No complaint about length was encountered earlier than the sixtieth item, but there were a few thereafter, suggesting that the seventy-four-item length represented something of a maximum if sustained interest were to be obtained from a majority of respondents unused to such procedures.

IV. FINAL ANALYSIS OF WHOLE INVENTORY

For the purpose of final selection of items to represent emotional maladjustment and sociability respectively, the writer waited until it was possible to accumulate an adequate sample. This was finally made up as shown in Table 1.

Table 1. *Composition of main development sample*

Omnibus conductor applicants,† accepted for employment, 'helping research'	100
Army (National Service men)	118
University students	19
Police drivers and instructors	14
Total, 'normals', research conditions	251
Omnibus conductor applicants,† inventory forming part of medical dept. examination before acceptance	100
Total, 'normals', varied conditions	351
Neurotic in-patients, mixed diagnoses†	27
Total	378

† Not part of the previous studies—all fresh samples.

* All the early development work on the second part of the inventory was carried out by the writer's colleague, Dr Peter Venables, who was also responsible for supervising the first-stage item analysis described in the next section of this report.

For this sample of 378 subjects, two sets of item analyses were carried out. These involved all 110 items (seventy-four plus thirty-six). The 378 sets of responses were placed in order according to total score for emotional maladjustment, and then divided into lower, middle and upper thirds. A count was then made of the number of times each item had been answered in the significant direction within each of these thirds. An item making a useful contribution to the total score would be expected to show a small total in the lower third and a large total in the upper third. If it were also discriminating in a perfectly linear fashion, the magnitude of the total for the middle third would be exactly intermediate, e.g. 20–40–60, 40–65–90. Using the two-fold requirement of linear discrimination and highly significant contribution to the original total score, twelve items were selected as a new total score to serve as the basis for a fresh item analysis.

A similar first-stage analysis was also carried out, again using all 110 items, in respect of the original thirty-six-item total score for sociability. In the selection of items for the second-stage analysis, a further requirement was added, namely, that an item selected to represent sociability should have made little or no contribution to the original total score for emotional maladjustment. On the basis of all three criteria (usefulness, linearity, independence), twelve items were selected as a new total score for sociability in the second analysis.

At the conclusion of the second-item analysis using the twelve-item total score for emotional maladjustment, it was found possible to select twenty items in all as a final total score. This twenty-item total score was then used as the criterion of emotional maladjustment when attempting to find further useful, linear and independent items to measure sociability. The stringency of this three-fold criterion, however, resulted in failure to produce a greater number of items, and the final total score for sociability had consequently to be based on twelve items.

To recapitulate: the twenty and twelve items finally chosen to represent emotional maladjustment and sociability respectively were those which had (*a*) shown a statistically significant contribution to the appropriate total score, (*b*) made this contribution in as linear a fashion as possible, and (*c*) as far as possible were *not* contributing significantly to the total score for emotional maladjustment when they were supposed to be measuring sociability.

V. Description and evaluation of the present two-part measure

This two-part personality measure consists of 110 statements which can be answered as 'True' or 'Not true' self-descriptions by the subject. The first part, seventy-four items, contains twenty items which are used as the basis for a score of emotional maladjustment. Apart from eight items of the MMPI 'Lie Scale', which can be used as a separate score if desired, most of the remaining items are buffers which cause the section as a whole to appear as a 'Health Inventory'. Average response time for this section is about 12 min. The odd-even self-correlation of the twenty-item score for emotional maladjustment, corrected to full length, is 0·81 ($N = 378$). Each item discriminates significantly at the 5% level, or better, between lower, middle and upper thirds of total score, when χ^2 on observed differences from chance expectancy is used.

As three-quarters of these items discriminate more sharply between middle and upper thirds than between lower and middle thirds of total score, despite efforts to find strictly

linear items, the distribution of scores shown in Fig. 1 is not surprising. It will be observed that scores up to 9 are distributed in a roughly normal fashion, and that there is a 'tail' from 10 upwards. One standard deviation above the mean for 251 'normals' falls at 9·03, one standard deviation below the mean for twenty-seven hospitalized neurotics falls at 9·48. It seems justifiable to regard scores of 10 and above as 'probably maladjusted'; scores from 0 to 7 as 'probably well adjusted', and to reserve scores 8 and 9 as 'doubtful'. The use of these groupings on the 251 'normal' respondents would divide them as to 13% 'probably maladjusted', 13% 'doubtful', and 74% 'probably well adjusted'.

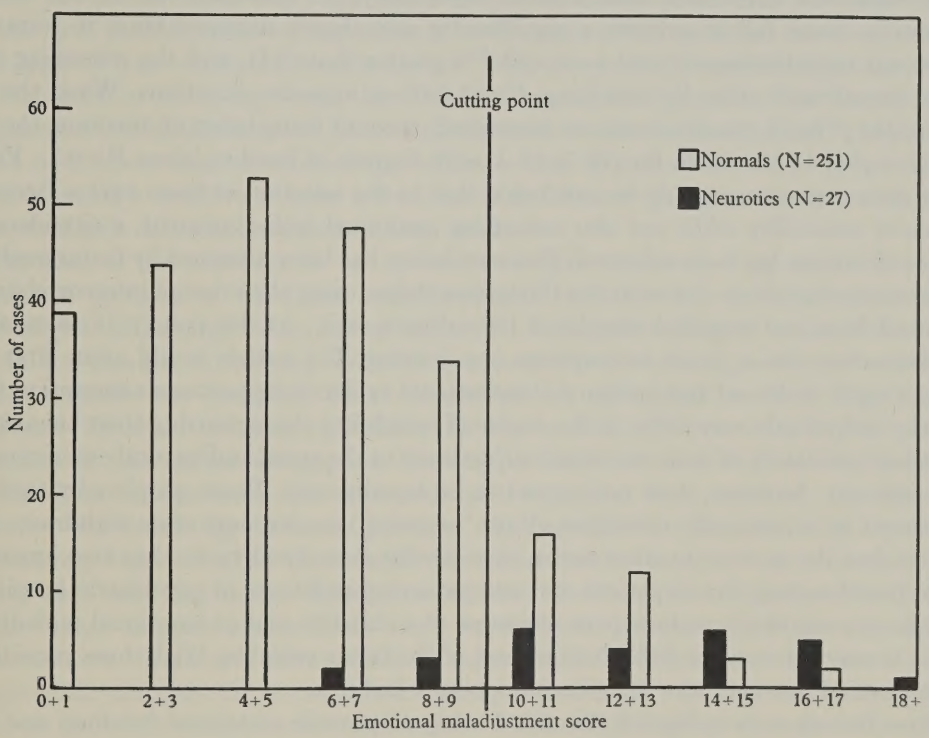


Fig. 1

For a fresh group ($N=100$) available as a check, the comparable figures are 13, 19 and 68%. In deciding to regard scores 0–7 inclusive as justifying a classification of 'probably well adjusted', the writer does not overlook the comments of Super (1942) and of Landis & Katz (1934) on the Bernreuter Inventory. These writers expressed doubts concerning the inferences to be drawn from a low score, with which the present investigator concurs. It is indeed unsafe to assume that a respondent obtaining a score lower than 8 on the present measure *is* well adjusted; all that can be said is that in the light of the available evidence such a respondent is likely to be well adjusted, and that the probability that this is so is reasonably high.

In view of data published elsewhere (Heron, 1956) it is essential to stress that these cutting-points and percentages cannot be considered to hold if the personality measure is used under conditions other than those usually obtaining in a research situation, and *certainly not under those obtaining in employment selection.* In brief, this is because the

evidence referred to makes it clear that selection conditions result in a general downward shift of the score distribution.

The second part of the measure consists of thirty-six items, twelve of which are used as the basis for a score of sociability. Average response time for this section is about 8 min. The odd-even self-correlation of this twelve-item score, corrected to full length, is 0·74 ($N = 378$), which is surprisingly high unless one pauses to consider the probably beneficial consequences of rigorous item analysis. Each item discriminates significantly at the 5% level or better between lower, middle and upper thirds of total score, when use is made of χ^2 on observed differences from chance expectancy. In addition, ten of the twelve sociability items fail to achieve a significantly non-chance discrimination in terms of emotional maladjustment total score (all P's greater than 0·1), and the remaining two items cancel each other by obtaining P's of 0·05 in opposite directions. When the independent χ^2 for the twelve items are summated, account being taken of direction, the net additive χ^2 is 17·55, which for (12–1) (3–1) = 22 degrees of freedom gives $P > 0{\cdot}7$. From these data it may reasonably be concluded that in the selection of these twelve items to represent sociability while not also measuring emotional maladjustment, a considerable degree of success has been achieved. This conclusion has been sustained by factor analysis of the intercorrelations between the thirty-two items, using the average intercorrelations obtained from two stratified samples of 100 subjects each. At this point it is essential to be clear about the *a priori* assumptions one is using. The author would argue that the great weight of clinical and ordinary observational evidence supports an assumption that healthy individuals may differ in the degree of sociability characterizing their behaviour, quite independently of their emotional adjustment to the social and natural environment. The contrary, however, does not appear to be equally true. Those people admitted for treatment as 'emotionally maladjusted' (or 'neurotic') are far more often withdrawn and asocial than the reverse; in other terms, there are far more dysthymics than true hysterics to be found among the in-patient and out-patient populations of psychiatric hospitals. If then one essays to produce pure measures of sociability and of emotional maladjustment, it may prove more difficult in respect of the latter variable. With these considerations in mind one may now examine the present findings.

Three factors were extracted (the third solely to provide rotational freedom) and the two expected factors of 'emotional maladjustment' and 'sociability' were clearly identifiable after two straightforward orthogonal rotations. The amounts contributed to the total variance accounted for by each of the factors can be apportioned between the twelve 'sociability' and the twenty 'emotional maladjustment' items respectively. When a correction is made to equate for number of items (at sixteen, the mean of twelve and twenty) it is found that the 'sociability' *items* only contribute 11% of the variance accounted for by the 'emotional maladjustment' *factor*. Conversely, however, the 'emotional maladjustment' items contribute 21% of the variance accounted for by the 'sociability' *factor*. From this it is clear that overlap between the two factors is mainly due to the impurity of items chosen to measure 'emotional maladjustment'. If one feels justified in assuming *a priori* that emotional maladjustment does not, on our pattern of culture, involve any element of withdrawal behaviour such as might be termed 'unsociableness', then a clear course of action is open. One need only discard items which measure both factors until one is left with as pure a group of items for emotional maladjustment as one has already obtained for sociability. However, examination of the raw

data concerning the twenty items used in the present study to represent emotional maladjustment raises the very problem anticipated in the theoretical discussion above. The items which one would discard in the interests of factorial purity tend to be those which most sharply discriminated between the validation groups of hospitalized neurotics and of apparently healthy people. One is therefore left wondering what a pure measure of 'emotional maladjustment' would represent when shorn of these items. By this time it might be safer to term the measure '*X*'. However, it could reasonably be argued that for practical purposes this would not matter, since to locate the emotionally maladjusted in any population under study one would simply use the bivariate distribution provided by the two 'pure' scores, and regard those in the appropriate quadrant (low 'sociability' plus high '*X*') as 'emotionally maladjusted'. Whether such a pragmatic approach would be theoretically satisfying would then remain a matter for debate and further research. The present instrument does not, however, lend itself to such treatment, which would, quite apart from theoretical considerations, require a very large pool of items. Use of such a large inventory would itself raise a fresh set of problems in the area of response motivation, to which reference has already been made.

The plan of development for the present two-part inventory required a valid, self-consistent measure of emotional maladjustment (whether this involved elements of pathological withdrawal behaviour or not) and a valid, self-consistent measure of *normal* sociability which could be used successfully as a criterion when no question of emotional maladjustment was involved. These objectives are considered to have been largely achieved. External evidence of validity for the sociability score must depend on acceptance of the earlier assumption relating 'introversion-extraversion' to 'anxiety-hysteria', but in this connexion some additional data have been obtained. An objective measure discriminating between patients diagnosed as 'dysthymic' and 'hysteric' respectively was evolved by Venables (1955). It consisted of a motor task carried out first under 'easy' and then under 'difficult' conditions. He found that hysterics showed little or no change in performance from 'easy' to 'difficult', whereas dysthymics tended to become more 'overactive' (Davis, 1948) under 'difficult' conditions. For twenty-four hospitalized neurotics the tetrachoric correlation was 0·64 ($P = 0{\cdot}001$) between Venables's measure and the twelve-item sociability score when a 13/11 split was made at the point of minimum misclassification.

Table 2 lists for information and interest the items used as a basis for the present total scores for emotional maladjustment and sociability, but it is strongly emphasized that no evidence is yet available which would justify extending the characteristics of these scores here reported to their use in isolation from the remaining buffer items. Psychologists wishing to use the two-part measure are invited to approach the writer for a copy of the full inventory and accompanying instructions.

Sources quoted are those from which the items were actually taken, but it is realized that many items also appear in slightly different form in other earlier inventories.

VI. Alternative methods of administration

In a paper read before a meeting of the British Psychological Society (Heron, 1953), the writer described a new method of administering inventories of this type. At the close of that paper, considerable emphasis was laid on the fact that no evidence was then available

Table 2. *List of items finally selected as basis for scoring*

First Part (emotional maladjustment)	S.R.	S.
3. I sweat very easily even on cool days	*T*	*w*
5. I am troubled by sick headaches	*T*	*u*
6. My feelings are easily hurt	*T*	*m*
7. Sometimes when I am not feeling well I am cross	*T*	*l*
9. I consider myself rather a nervous person	*T*	*m*
15. I never have nightmares	*F*	*m*
20. I wake up fresh and rested most mornings	*F*	*w*
23. My eyesight is sometimes blurred	*T*	*p*
25. I worry over possible misfortunes	*T*	*m*
27. It is difficult to irritate me	*F*	*m*
33. My bowels are not opened every day	*T*	*p*
44. My physique is excellent	*F*	*p*
46. Once in a while I think of things too bad to talk about	*T*	*l*
47. I often feel self-conscious in the presence of superiors	*T*	*m*
50. I feel that very little works out the way it should in life	*T*	*c*
55. I have no particular physical or health problems	*F*	*u*
60. I have periods of such great restlessness that I cannot sit long in a chair	*T*	*w*
61. My hands often feel damp and clammy	*T*	*u*
69. I catch cold frequently	*T*	*p*
72. I often jump to conclusions	*T*	*r*
Second Part (sociability)		
79. The presence of the opposite sex makes me feel shy	*T*	—
82. What others think of my actions does not bother me much	*F*	—
83. I sometimes treat my work as if it were a matter of life and death	*T*	—
84. I like to play practical jokes on people	*F*	—
86. I find it easy to act naturally at a party	*F*	—
93. I like to have time to be alone with my thoughts	*T*	—
97. I like to have a lot of social engagements	*F*	—
101. I tend to be overconscientious	*T*	—
102. I am usually a good mixer	*F*	—
103. I think of myself as a lively sort of person	*F*	—
104. When I am bored I feel like stirring up some excitement	*F*	—
109. At a gay party I can usually let myself go and have a thoroughly good time	*F*	—

S.R. significant response; S. source.

Key: *w*, Winne's MMPI study; *u*, U.S. Army; *m*, Maudsley Medical Questionnaire; *l*, Lie scale, MMPI; *p*, physical health buffer item originally; *c*, Cattell; *r*, rhathymia item transferred; *T*, response significant if marked 'True'; *F*, if marked 'Not true' (=False).

Note. The 'significant responses' shown in the second part are in the '*un*sociable' direction for convenience in scoring.

to show what difference if any might be made as a result of using the new method as compared with a conventional pencil-and-paper approach. It is now possible to state that so far as means of total score are concerned, no significant differences arise between the two methods. The new method of administration, which employs a sealed box and cards, has, however, been described by several users as generally more acceptable than the usual pencil-and-paper presentation, possibly because it provides a certain amount of response motivation.

Only those who have tried to gather large-scale data will appreciate how greatly the writer has depended on the help of others. His first expression of appreciation, none the less real because perforce anonymous, must go to the 1696 men and women who have during the past four years completed responses to this inventory, at various stages of its development, and by one or other of the methods of administration.

Prof. H. J. Eysenck assisted with advice and facilities when both were especially needed. Mr A. E. Maxwell, M.A., frequently gave valuable statistical advice during the final stage, including the specific suggestion which led to the first-stage item analysis procedure.

Prof. Rex Knight, Miss Betty Fraser, Mrs Lorna Simpson, and students at the University of Aberdeen combined forces to ensure that responses from university students, a category usually responsible for all the data concerning measures such as these, were not on this occasion wholly excluded. To Ian Haldane, M.A., of the Army Operational Research Group, and Inspector F. R. Priestley of the Essex County Police goes appreciation for their cooperation in extending the range of the main sample.

The co-operation of workers in electrical engineering factories was obtained through the courtesy of management and trade union at Ferranti Ltd; of new entrants as omnibus conductors in Manchester, by the help of the General Manager, A. F. Neal, B.Sc. (Eng.), Mr J. F. J. Webb, Personnel Manager, Mr E. Darby, Chief Instructor of the Training School, and Councillor Thomas, Transport and General Workers Union. Similar assistance in London was obtained through the good offices of the Chief Medical Officer, Dr L. G. Norman, and of the Recruiting, Training and Education Officer, Mr K. R. Thomas, and their staffs at the Central Recruiting Office.

The development of this criterion formed part of a long-term programme of research into the association between psychological and occupational handicap, under the general direction of Prof. Aubrey Lewis, M.D., F.R.C.P., whose comments and criticisms have been sincerely appreciated. The criticism and assistance of Unit colleagues is also gratefully acknowledged.

REFERENCES

DAVIS, RUSSELL (1948). *Pilot Error*. London: H.M. Stationery Office.

EYSENCK, H. J. (1952). *The Scientific Study of Personality*. London: Routledge and Kegan Paul, Ltd.

EYSENCK, H. J. (1953). *The Structure of Human Personality*. London: Methuen and Co. Ltd.

HERON, ALASTAIR (1953). Old Wine in a New Bottle: a new method of presenting a psychiatric questionnaire. Abstract. *Quart. Bull. Brit. Psychol. Soc.* no. 20, p. 15.

HERON, ALASTAIR (1955*a*). Personality and occupational adjustment: a cross-validation study. *Canad. J. Psychol.* **9**, 15–20.

HERON, ALASTAIR (1955*b*). The objective assessment of personality among female unskilled workers. *Educ. Psychol. Meas.* **15**, 117–26.

HERON, ALASTAIR (1956). The effect of real-life motivation on questionnaire response. *J. Appl. Psychol.* **40**, 65–68.

HILDEBRAND, H. P. (1953). A factorial study of introversion-extraversion by means of objective tests. Unpubl. Ph.D. thesis, University of London.

LANDIS, C. & KATZ, S. E. (1934). The validity of certain questionnaires which purport to measure neurotic tendencies. *J. Appl. Psychol.* **18**, 343–86.

SUPER, D. E. (1942). The Bernreuter Personality Inventory. *Psychol. Bull.* **39**, 94–125.

VENABLES, P. H. (1953). A study of motor and autonomic responses to experimentally induced stress in an industrial setting. Unpubl. Ph.D. thesis, University of London.

VENABLES, P. H. (1955). Changes in motor response with increase and decrease in task difficulty in normal industrial and psychiatric patient subjects. *Brit. J. Psychol.* **46**, 101–110.

WINNE, J. F. (1951). A scale of neuroticism: an adaptation of the Minnesota Multiphasic Personality Inventory. *J. Clin. Psychol.* **7**, 117–22.

(*Manuscript received* 21 *April* 1955)

SOME EXPERIMENTS ON PAVLOV'S THEORY OF SUGGESTION*

By J. G. INGHAM

Regional Neuropsychiatric Research Centre, Whitchurch Hospital, Cardiff and Morgannwg Hospital, Bridgend

Pavlov's theory implies that some degree of inhibition should be present following a period of suggestion and that the amount of inhibition will to some extent determine the response to suggestion. Experiments were carried out to determine whether changes in auditory sensitivity following suggestion were in accordance with this theory. As predicted, suggestions of arm movement were followed by increased thresholds. A period of prose reading of the same duration and intensity produced the same effect but casual conversation did not. Correlations between suggestibility and threshold increment were, on the whole, positive but not large enough to justify a definite conclusion. The threshold increment was greater in subjects who had been taking sedative doses of sodium amytal for one week.

I. Introduction

A response to suggestion has often been said to be determined by two processes, usually antagonistic. Eysenck (1947), for instance, spoke of 'aptitude' or ideo-motor tendency and 'attitude' or the conscious inhibition of this tendency. Hull (1933) put forward a similar proposal. Pavlov (1927) explained suggestion in terms of his general theory of brain function. This led him to suggest one way in which the two processes might interact. According to him, suggestion causes a concentration of excitation in part of the cortex while other areas are inhibited by negative induction. This diminishes the effect of influences which would otherwise tend to counteract the suggestion.

From Pavlov's theory one can predict first that some degree of inhibition will be present following a period of suggestion, and secondly that the amount of inhibition will to some extent determine the response to suggestion. There should in fact be a positive correlation between response and inhibition. The object of the first experiment was to test both these predictions.

II. Experiment 1

(i) *Method*

Monaural auditory thresholds were determined before and after a test of suggestibility. The horizontal arm-movement test (Ingham, 1955) was used, the suggestions lasting for 2·5 min. Changes in threshold were compared with those in a control group of different individuals for whom, instead of suggestion, a period of casual conversation intervened between the threshold measurements. Subjects wore headphones, and the stimulus for the hearing tests was a 400 cycles per sec. tone presented to the left ear. Thresholds were determined by the method of limits using series of increasing intensities. Three such determinations were made before the suggestion or control period and three after. Before this, however, six or more preliminary practice determinations were made. All these measurements were made in groups of three with a rest interval 3–4 min. between groups.

* This article is based on a paper read at the Annual Conference of the British Psychological Society in Durham, April 1955.

Within each group there was half a minute between the completion of one determination and the beginning of the next. The subliminal intensity of the tone at the start of the threshold determinations was varied throughout the practice groups but was otherwise held constant. Catch tests were introduced during the practice groups.

Suggestions were given by means of a tape recorder connected to the right headphone, the gain control being set to give comfortably audible speech reproduction. During the equivalent intervening time interval the control subjects kept their headphones in position but received no sound stimulus through them. They conversed with the experimenter who avoided topics likely to evoke emotional reactions.

The subjects were male adult neurotic in-patients selected as being co-operative and sufficiently intelligent to follow the instructions. Patients with a previously detected impairment of hearing were excluded. Seven subjects were allotted at random to each group and the experiment was repeated on five successive days for each subject.

(ii) *Results*

In Figs. 1 and 2 are the mean thresholds before and after the suggestion or control period. These are the results for the first 2 days only. All 5 days showed a similar pattern and, after Exp. 2, only two repetitions per subject were used.

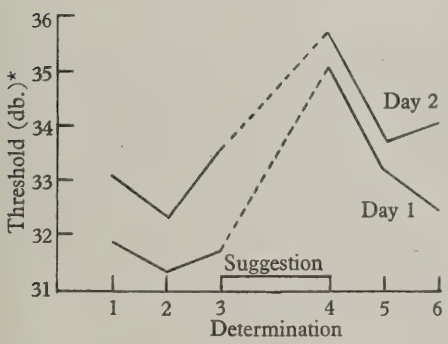


Fig. 1. Exp. 1 mean thresholds (experimental group). * Reference level = 0·001 mV. (r.m.s.) at headphone.

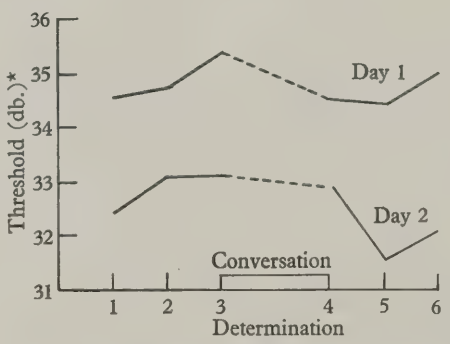


Fig. 2. Exp. 1 mean thresholds (control group). * See footnote to Fig. 1.

For each session and for each subject the arithmetic mean of determinations 1, 2 and 3 was subtracted from determination 4. This difference will be called the threshold increment. The mean threshold increments for the 2 days were +3·1 db. (experimental group) and −0·2 db. (control group). There was very little overlap between the two groups, and the probability of getting as large a difference in the predicted direction by chance is less than 0·005 ($t = 3·34$). There was no overlap between the groups when the mean threshold increments for the 5 days were compared. It appears, therefore, that the first prediction has been verified; there is an increase in auditory threshold following a period of suggestion.

Table 1 shows the rank correlations (τ) (Kendall, 1948) within the experimental group between suggestibility and threshold increment. They are consistent with the second prediction, that there should be a positive correlation between the two variates, but the levels of significance do not justify a definite conclusion.

Table 1. *Correlation between suggestibility and threshold increment* ($N=7$)

Day	r	P
1	+0·52	0·068
2	+0·39	(0·19)*
3	+0·35	0·19
4	−0·16	—
5	+0·39	0·15
Total	+0·33	0·19

* Approximately.

III. Experiment 2

(i) *Method*

The object was to repeat Exp. 1 using different control conditions. This time the control subjects listened to an auditory stimulus of the same duration and intensity as the suggestion. A descriptive passage from Hardy's novel *The Woodlanders* was used. In order to equate the intensities as far as possible it was necessary to make a new recording of the suggestions. There were other slight changes in technique aimed at reducing experimental errors. Subjects were selected from a similar group to those in Exp. 1. Ten were allotted at random to the experimental group and the remaining eleven to the control group.

(ii) *Results*

Average thresholds are shown in Figs. 3 and 4. Although the first post-suggestion average threshold was slightly higher than the three initial points on both graphs for experimental subjects, it was clear that the marked increments obtained previously were

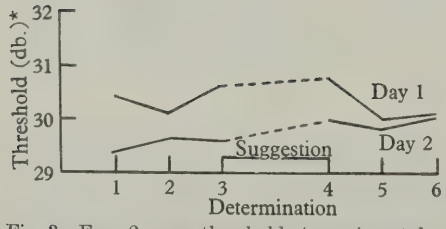


Fig. 3. Exp. 2 mean thresholds (experimental group). * See footnote to Fig. 1.

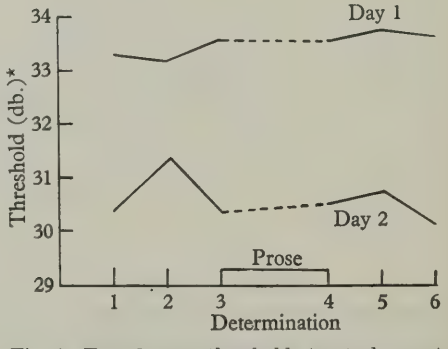


Fig. 4. Exp. 2 mean thresholds (control group). * See footnote to Fig. 1.

no longer occurring. There was no significant difference in threshold increment between experimental and control groups. There was again some slight evidence of a correlation between threshold increment and suggestibility (Table 2).

There was clearly a discrepancy between these results and those of the first experiment. The probability of both sets of threshold increments following suggestion being selected by chance from the same population was one in four hundred. It became necessary to ascertain the reason for this discrepancy.

IV. Experiment 3

The procedures used in the first two experiments were compared for differences which might account for the failure to produce the effect the second time. Two changes which had been made in Exp. 2 were the introduction of a new recording of the suggestions and the elimination of a click in the headphone which had previously followed the suggestion. To see whether either the old recording or the click were responsible for the effect in Exp. 1 a group of similar patients were tested once using the old recording without the click and once using the new recording with the click. Subjects were tested on successive days in counterbalanced order. In neither case was the result materially different from that of Exp. 2, and there was no significant difference between mean increments under the two conditions. Correlations between threshold increment and suggestibility ($\tau = -0{\cdot}1$ and $0{\cdot}0$) were insignificant.

Table 2. *Correlation between suggestibility and threshold increment* ($N = 10$)

Day	τ	P
1	+0·25	0·19
2	−0·51	—
3	+0·11	0·36
4	+0·64	0·0046
5	+0·33	0·108
Total	+0·33	0·108

V. Experiment 4

(i) *Method*

Another possibility emerged from an examination of the medical records of the subjects. It was found that most of the patients taking part in the first experiment were, and had been for some time, receiving fairly large doses of sedatives, sodium amytal in most cases. The subjects of the later experiments, however, were rarely sedated, and the few who were, had received considerably smaller doses. The object of Exp. 4 was to see whether this could account for the difference between the two sets of results. Exp. 2 was repeated using two groups of patients, one receiving sodium amytal in doses comparable with those of Exp. 1 and the other unsedated. The subjects, all neurotic patients, were allotted at random to the sedated or unsedated groups. Six of the sedated subjects received 9 grains per day given orally and the other nine received 4 grains per day by the same route. The medication had continued for 1 week prior to the experiment and the unsedated subjects had remained unsedated during the same period. Two experimental sessions per subject were undertaken on successive days, one under the experimental conditions (i.e. with suggestion intervening between the two sets of threshold measurements) and one under the control conditions (i.e. with a passage of prose intervening). These were the stimulus conditions used in Exp. 2. The order of the experimental and control conditions was alternated at random throughout the groups. Two hypotheses to be tested were (1) that suggestion would produce a greater threshold increment than the control conditions and (2) that the sedated group would show greater threshold increments than the unsedated group. There were fifteen sedated subjects and thirteen unsedated.

(ii) *Results*

The average thresholds for the control and suggestion conditions are shown in Fig. 5. There was no significant difference between the two sedated groups which have therefore been combined. One point that seems immediately apparent is that these results do *not* support the first hypothesis. There was no significant difference between suggestion and control conditions in either sedated or unsedated groups.

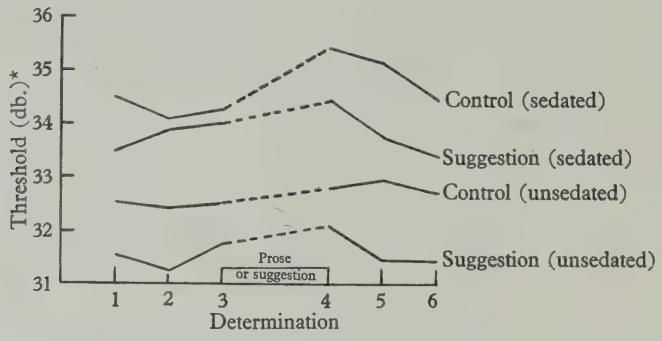


Fig. 5. Exp. 4 mean thresholds. * See footnote to Fig. 1.

Table 3. *Analysis of threshold increments in decibels (mean of 2 days)*

	N	Mean	*t*	*P*
Exp. 4: Sedated	15	+0·96	3·22	<0·005
Unsedated	13	+0·46	2·24	<0·025
Exp. 1: Sedated (controls)	6	−0·06	(1·77)*	(<0·05)*

* Refers to the difference between Exp. 1 sedated (controls) and Exp. 4 sedated.

Table 3, however, shows that there is a significant increase in threshold following the period of stimulation for both sedated and unsedated groups. The lower row shows that the mean increment for the sedated controls in Exp. 1 (with casual conversation during the control period) was slightly less than zero and was significantly less than that for the sedated group of Exp. 4.

Pavlov's theory specifies that suggestion should produce inhibition by negative induction. A difficulty arises, however, when we try to specify stimulus conditions under which the same phenomenon would *not* be expected to occur. According to Pavlov a stimulus exciting the cortex may cause one of two effects. Either the excitation is concentrated, in which case it induces inhibition in surrounding areas, decreasing the excitability of these areas, or, alternatively, the excitation irradiates, thereby increasing the excitability of surrounding areas. Both irradiation and concentration may occur at different stages during and following the action of a single stimulus. Unfortunately, he was unable to specify unambiguously the conditions under which one or the other would predominate. Because of this it is difficult to set up, with any conviction, a stimulus situation which would *not* be expected to produce negative induction.

When a passage of prose was chosen for the control situation, the implicit hypothesis was that the repetitive quality of the suggestion was the characteristic necessary to produce negative induction. The effect of such a stimulus was therefore compared with

another which was not repetitive but included a variety of ideas in a passage of descriptive writing. The results are not in accord with the hypothesis. An important problem for future research is to discover the feature or features responsible for producing the inhibitory effect in those situations where it occurs.

Let us now consider the second hypothesis, that the sedated group will show greater increments than the unsedated group. As shown in Table 3 the mean increment of the sedated group was 0·5 db. greater than that of the unsedated group. The probability of getting, by chance, as large a difference in this direction is between 0·10 and 0·05 using a t-test ($t = 1{\cdot}32$). Using a non-parametric method (Kendall, 1948) the probability is 0·06. As the difference was of borderline significance it was considered advisable to increase the size of the groups. It was possible to do this for the unsedated group by adding to it those subjects from the earlier experiments who had been tested under the same stimulus conditions (either suggestion or prose) and who were not receiving sedatives. The difference between the groups was now significant at the 5% level. The data, therefore, lend definite support to the hypothesis. It seems likely that the inhibitory effect is increased by sodium amytal under these conditions.

Does this account for the discrepancy between the results of Exp. 1, where the subjects were sedated, and those of Exps. 2 and 3 where they were not? In Exp. 1 six of the seven patients in the experimental group were taking sodium amytal at the time of the experiment. The mean threshold increment for these six was +3·55 db. The difference between this and the +0·96 db. which was the equivalent figure for the sedated subjects in Exp. 4 was significant ($t = 3{\cdot}50$ and $P < 0{\cdot}01$ using a two-tail test). It is clear that some factor is required to explain the results of the first experiment in addition to sedation. One such factor is the longer period of medication involved. The six sedated patients in Exp. 1 had been receiving sodium amytal as a part of their usual treatment for considerably longer than 1 week. Another possibility is that a modification of instructions made after Exp. 1 was sufficient to reduce the effect. Following the suggestion in Exp. 1 and preceding the first post-suggestion threshold measurement the investigator said 'listen again now please'. In later experiments the words '. . . tell me as soon as you can hear the sound' were added. It is possible that attention was more effectively concentrated on the tone by the second instruction. There may be other differences between the circumstances of the two experiments of which we are not aware.

In Exp. 4 there was no evidence of a correlation between suggestibility and threshold increment. A new apparatus was being used to measure arm movements, and an unanticipated source of error was discovered in the course of the investigation. It was felt that less reliance could be placed upon these suggestibility measurements than in Exps. 1 and 2, where some evidence of a positive correlation was found.

VI. Conclusions

The object of these experiments was to test two predictions from Pavlov's theory of suggestion. The main findings were:

(1) As predicted by the theory, a period of suggestion, given under the conditions of the arm-movement test, was followed by a reduction in auditory sensitivity.

(2) A period of prose reading, of the same duration and intensity and under the same conditions, produced the same effect.

(3) A period of casual conversation of the same duration produced no such effect.

(4) The second prediction (of a positive correlation between suggestibility and threshold increment) was neither confirmed nor disproved, but the evidence, as far as it goes, is slightly more for it than against it.

(5) The threshold increment was greater in subjects who had been taking sedative doses of sodium amytal for 1 week.

I am indebted to Mr M. B. Shapiro for the discussions which led to the formulation of the first experiment. Drs T. J. Hennelly, J. M. White, D. Lewis and C. J. Morgan kindly allowed me to use patients under their supervision as subjects, and Dr Morgan supervised the administration of drugs in Exp. 4. I also wish to thank the patients themselves for their willing co-operation. I am grateful to Dr Derek Richter, Director of Research, for his interest and advice.

References

Eysenck, H. J. (1947). *Dimensions of Personality.* London: Kegan Paul.
Hull, C. L. (1933). *Hypnosis and Suggestibility.* New York: Appleton-Century-Crofts.
Ingham, J. G. (1955). Psychoneurosis and suggestibility. *J. Abnorm. (Soc.) Psychol.* **51**, 600–3.
Kendall, M. G. (1948). *Rank Correlation Methods.* London: Griffin.
Pavlov, I. P. (1927). *Conditioned Reflexes.* Oxford University Press.

(*Manuscript received* 28 *May* 1955)

THE EFFECTS OF PROPAGANDA ON RECALL, RECOGNITION AND BEHAVIOUR

II. THE CONDITIONS WHICH DETERMINE THE RESPONSE TO PROPAGANDA

By EUNICE BELBIN

Formerly of the Psychological Laboratory, University of Cambridge

A series of experiments were conducted in which Road Safety Posters were displayed in a waiting room. Their effect was measured by: (*a*) the amount of information from the posters which was applied to the interpretation of photographs showing faulty traffic situations (*Use* scores); and (*b*) the amount of information contained on the posters which could be recalled (*Recall* scores). An attempt was made to investigate the reasons why on some occasions or with some subjects the information was recalled more than it was used, and why on other occasions or by other subjects it was used more than it was recalled.

Four types of propaganda were displayed. Each had a significant effect on the use scores, but there was no significant difference between the effects of these four types on either use or recall. However, the previous experience of the subjects was shown to have a considerable effect on the results. The over-sixties, although comparatively bad at recalling the propaganda, had use scores which compared favourably with those of the younger subjects. The subjects who had car-driving experience used the information from the posters to a significantly greater extent than the non-drivers. On the other hand, the drivers did not recall the propaganda any better than the non-drivers. There was also a tendency for the naval rating subjects to use more of the information than they recalled and for the students to recall as much or more than they used. When the tests were applied up to a period of 14 days after the display of propaganda there was a widening disparity between use and recall scores. At 14 days subjects still showed a slight tendency to use the information although they were not able to remember it.

An attempt is made to explain these results by suggesting that when a subject is presented with an incidental learning situation, the meaningfulness he extracts from the material and his method of organizing it depends on his previous learned responses to similar material. Where his previous experience has entailed incorporation of the information into his activity, then the new material to be learned is likely to have meaning in relation to its practical application, a response which favours the retention of the information for use in other tasks, but not necessarily for recall. If his previous experience has been learning by memorizing, his response to the new material will be organized accordingly; he is likely to learn in order to recall.

I. Introduction

In some previous experiments which have been conducted by the writer (Belbin, 1956) it became evident that the information which had been learned from propaganda material could become apparent in two distinct forms, namely, (*a*) in recall or recognition, and (*b*) in subsequent behaviour.* For example, when Road Safety Propaganda was presented to children in their Cinema Clubs it was found that although the propaganda had no measurable effect on their behaviour, they could nevertheless describe and recognize it quite well. Children who were best able to describe the propaganda material were not always those who behaved correctly on the roads. Again, when propaganda was used in relation to performance on a laboratory task it was found that some subjects were

* Throughout this paper the term behaviour has been used to denote non-verbal action.

unable to describe it although it affected their behaviour. Others who could describe or recognize it were not influenced by it.

These results are in conflict with the assumption which has hitherto been made that propaganda is likely to affect behaviour in proportion to the extent to which it can be recognized or recalled. It was not clear, however, from these experiments what caused the different effects to be more marked in some subjects than in others. The following series of experiments (reported more fully elsewhere, see Belbin, 1953) were therefore designed to examine this problem. In these experiments we have tried to discover the reasons why the information contained in propaganda should on some occasions be remembered but not used and at other times be used but not remembered. We have examined the extent to which these different responses might be a function of: (1) the type of propaganda; (2) the characteristics or experience of the subject; and (3) the length of time between display of the propaganda and the measure of its effect.

II. Experimental procedure

Six Road Safety Propaganda Posters were displayed in a waiting room. Subjects were left in this room without instruction for 3 min. Afterwards they were asked to point out, in some photographs of local traffic scenes, any faults they could find in relation to Road Safety. Each of six photographs (test photos) was directly related to one of the six posters, the purpose of the experiment being to find the effect of the posters on the interpretation of the photographs. A further eighteen photographs (non-test photos) which were not related to the posters were shown to the subjects along with the test photos. After the photograph test, the subjects were asked to describe the posters which they had seen in the waiting room. A control group of subjects was asked to interpret the same set of photographs without propaganda having been displayed previously. The experimental design was therefore as follows:

	Experimental group	Control group
1.	Display of posters (3 min.)	—
2.	Photo test (18 non-test photos, 6 test photos)	Photo test (18 non-test photos, 6 test photos)
3.	Recall test (description of photos)	—

It was thus possible to compare under various conditions: (*a*) the extent to which the information contained in the posters was applied to the photo test (use scores); and (*b*) the extent to which the propaganda was remembered (recall scores). Subjects scored one mark on use scores for each test photo when the reply was in accordance with the propaganda. For recall scores one mark was scored when it was clear from the description that the poster had been correctly interpreted.

Four different types of propaganda were shown: (*a*) pictorial; (*b*) horror; (*c*) verbal positive; and (*d*) verbal negative. There were six posters for each type, the slogans of these six for types (*a*), (*b*) and (*c*) being as follows: (1) Dog-owners! Keep your dog on a lead. (2) Cyclists! Keep well away from stationary cars—the door may suddenly open. (3) Pedestrians! Give way to traffic on crossings when the policeman signals it on. (4) Cyclists! Only very small children may be carried as passengers. (5) Pedestrians! Always face the oncoming traffic. (6) Cyclists! Strap your parcels on. A hanging package may catch in the wheel.

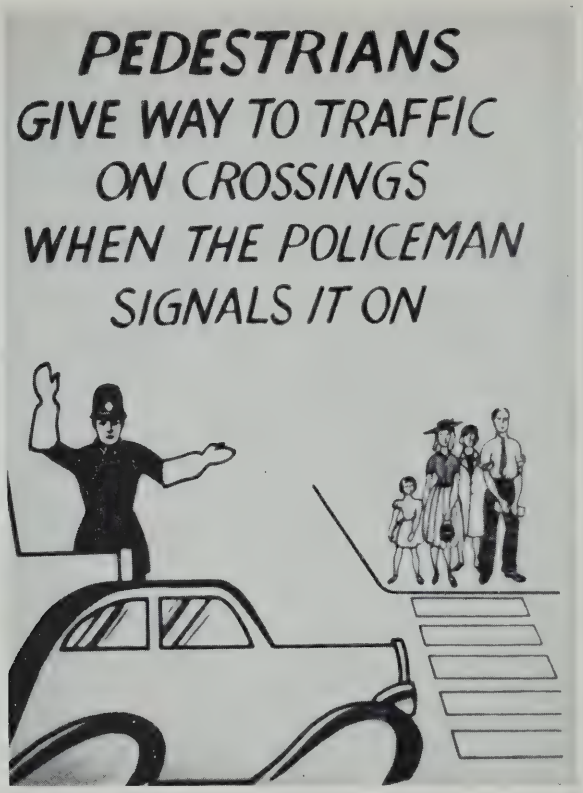


Fig. 1. Copy of one of the posters which was used in the experiment. (Pictorial Propaganda Number 3)

Fig. 2. The Test Photo which was related to the Poster shown in fig. 1.

Type (*c*) merely carried these slogans; type (*a*) also showed pictures of the correct behaviour to be used; type (*b*) displayed accidents which were resultant upon incorrect behaviour; and type (*d*) contained the same messages but in negative form, e.g. no. 5 was 'Pedestrians! Never walk with your back to the traffic.' Pl. 1, fig. 1, shows one of the pictorial posters used and Pl. 1, fig. 2 shows the corresponding photograph which was included in the subsequent test.

The comparative effects of these four different types of propaganda were observed both in use and recall scores; (*a*) with a group of fifty-seven naval ratings aged 18–22 years; and (*b*) with a group of fifty-five university students aged 18–22 years. The effect of the pictorial propaganda was also assessed on: (*c*) a group of men and women aged 65–75 years who were drawn from the Over-Sixties Clubs ($N = 21$). In all three cases part of the group acted as a Control and did not see the propaganda, but merely interpreted the photographs. Finally, the test was presented with different time intervals between the display of the propaganda and the photo test to: (*d*) six groups each comprising five men and five women aged 18–22 years. The men were university students; the women were from a Teachers' Training College.

III. RESULTS

(1) *The effect of type of propaganda*

The mean scores for the groups of students and naval ratings after each of the four types of propaganda are shown in Table 1, together with the values of $S\sqrt{(\text{var } S)}$ (rank correlation, Whitfield's method (1947)) when the use scores of each group are compared with those of the control group. It may be noted that: (*a*) All types of propaganda significantly affected the use scores on the photo test. This was true for both students and for naval ratings. (*b*) The propaganda had no significantly greater or less effect on the use scores of the students than on those of the naval ratings. (*c*) The greatest difference in use scores within the groups was between pictorial and horror propaganda for the naval ratings, pictorial being the more effective. The difference was not significant ($S\sqrt{(\text{var } S)} = 0{\cdot}75$). (*d*) There was no significant difference between the amounts recalled after different types of propaganda. And (*e*) when the differences between use and recall scores (use minus recall) were considered for each type of propaganda, no significant difference was found between any two types.

(2) *The characteristics or experience of the subject*

Although there was no significant difference between the mean scores of the naval ratings and of the students after the different types of propaganda, there was, nevertheless, considerable individual variation within the groups. Information concerning the naval ratings which had been recorded during the experiment was analysed to see whether it could account for any of these differences.

(*a*) *Car driving.* The scores of the naval-rating drivers (or motor cyclists) were compared with those of the naval rating non-drivers. For every type of propaganda the drivers had higher mean use scores than the non-drivers. For example, after Pictorial Propaganda the mean use score of drivers ($N = 6$) was 5·0 as compared with 2·6 for non-drivers ($N = 8$). There was a significant difference between the use scores of drivers ($N = 15$) and those of non-drivers ($N = 30$) when all the scores for the different types of propaganda were taken

Table 1. *Use and recall scores after different types of propaganda*

	Use				Recall
	No. in group	Mean score (6 photos)	$S^*/\sqrt{(\text{var } S)}$	Sig. at* 5% level (i.e. >1·96)	Mean score (6 posters)
Naval ratings:					
Control group	12	1·4	—	—	—
Pictorial	14	3·6	3·2	Yes	2·6
Horror	10	2·9	2·4	Yes	2·6
Verbal positive	11	3·1	2·9	Yes	1·9
Verbal negative	10	3·2	3·2	Yes	2·9
Students:					
Control group	12	1·8	—	—	—
Pictorial	11	3·6	2·9	Yes	3·1
Horror	11	3·6	2·4	Yes	2·2
Verbal positive	10	3·5	2·5	Yes	2·8
Verbal negative	11	3·0	2·1	Yes	1·9

* When the scores after each type of propaganda are compared with the Control group scores.

together. Yet no propaganda was specifically addressed to drivers; two posters seemed quite inappropriate to them. There was no evidence to suggest that these results were determined or biased by any particular poster or posters.

This difference in the effect of propaganda on drivers as compared with non-drivers *cannot* be explained by the fact that car drivers were better at interpreting road safety photographs in any case. The mean scores on the eighteen non-test photos were as follows: drivers ($N=10$) 15·1; and non-drivers ($N=10$) 15·3. Thus without propaganda naval-rating drivers were slightly worse than non-drivers; with propaganda, they were significantly better. Nor can it be explained by the fact that drivers were more 'intelligent' than non-drivers, for there was no significant relationship between I.Q. and car-driving. Further, this difference cannot be due to the fact that car drivers had read the Highway Code. If this had accounted for the difference in performance, one would expect the scores on all photos to be higher for car drivers than for non-drivers. (This was not so, as shown above.) Secondly, two of the posters were intentional deviations from the Code, as follows:

Highway code instructions	Poster slogans
"...do not carry a passenger on your cycle unless it is constructed or adapted for the purpose	Only very small children may be carried as passengers
"...Motorists—before opening any door of a vehicle make sure that...and that you will not inconvenience anyone on the road or footpath...'	Cyclists—keep well away from stationary cars. The door may suddenly open

It might have been expected that, if the Highway Code directly influenced the results, the drivers' responses to the photos would have been in accordance with the instructions in it rather than in accordance with the poster instructions. In actual fact 87% of the drivers' responses were in accordance with the 'carrying passengers' poster (compared with 50% of those of the non-drivers) and 67% of the drivers' responses were in accordance with the 'opening car doors' poster (compared with 47% of the non-drivers).

The numbers in the driver groups were too small to enable a comparison to be made between the relative effects of different types of propaganda on drivers and non-drivers. The numbers in the non-driving groups, however, were slightly larger. From a consideration of these scores type of propaganda seemed to be unimportant. It was not possible to make a similar analysis for the women students as too few of them drove cars.

Let us now consider the recall scores. A very different relationship between the response of drivers and of non-drivers is apparent. For example, after Pictorial Propaganda, whereas the use scores had been 5·0 for drivers ($N=6$) and 2·6 for non-drivers ($N=8$), the mean recall scores for these subjects were 3·0 and 2·4 respectively. Thus while there was a significant difference between the use scores for the two groups, no significant difference existed between their recall scores $S\sqrt{(\text{var}S)}=0{\cdot}22$. For five drivers with recall scores of 3 or 4 (3, 3, 3, 3, 4) the use scores were 6, 6, 6, 6, 4. On the other hand, for six non-drivers with recall scores of 3 or 4 (3, 3, 3, 3, 3, 4), the use scores were 3, 3, 3, 1, 4, 2. Similar effects could be noted after the display of Horror Propaganda.

It was not possible to consider whether cycling experience had similar effects to those of car driving. Nearly every subject was a cyclist.

(*b*) *Accidents.* It was thought that subjects who had had (or seen) bad accidents might have been more responsive to the propaganda or might have remembered it better. However, there was a slight tendency for those who had not had accidents to score more both on recall and on use than those who had had accidents. Use minus recall differed very little for the two groups.

(*c*) *Intelligence quotient.* All the naval-rating subjects were given an A.H.4 test. There was a significant positive correlation (at 5% level) after Pictorial Propaganda between A.H.4 scores and use scores and between A.H.4 scores and recall scores, but not after other types of propaganda. There was no positive relationship between 'intelligence' and general ability to interpret the photos. When the A.H.4 scores were measured against non-test scores $r=-0{\cdot}05$. With the relatively homogeneous group of subjects, whose A.H.4 scores fell mainly in the B category of the industrial norms, any relationship that might have existed between intelligence quotient and scores after the propaganda might easily have been masked. For example, eight out of ten subjects in the Horror group were in category B of the industrial norms. There was, however, a negative correlation ($r=-0{\cdot}58$) between A.H.4 score and number of misinterpretations of the meaning of the Horror Propaganda (apparent from the descriptions in the recall test), i.e. the less 'intelligent' were more likely to misinterpret the posters. Similarly, naval ratings misinterpreted the Horror posters one and a half times as often as the students. (There were only a few misinterpretations with the other types of propaganda.)

(*d*) *Sex differences.* It was not possible to consider sex differences with the foregoing data. For one reason most of the women students were non-drivers. However, a group of men students (aged 18–22) were tested after seeing the Pictorial Propaganda. They had a mean use score of 3·3 (when 10/11 were drivers). The woman students had a mean use score of 3·6 (when 10/11 were non-drivers). It is possible, therefore, that the women students might have been more influenced by the Pictorial Propaganda than the men students if the groups could have been equated in respect of driving experience. The mean recall scores for the two groups were: women, 3·1; men, 3·2.

(*e*) *Age differences.* Since there is a marked increase in the national pedestrian accident rate of people over the age of 65, the effect of propaganda on this group is of great practical importance. As most of the subjects, drawn from the Over-Sixties Club, were over 70, they were unlikely to have had recent experience in riding bicycles or driving cars, and it was therefore thought reasonable to consider only the effect of the three posters most relevant to the group (i.e. posters 1, 3 and 5). In addition, a number of the subjects were short-sighted. A smaller waiting room was therefore used where they were able to sit

within a few feet of the display posters. When their mean scores were doubled (i.e. to enable a rough comparison to be made with the groups who had seen all six posters) the Control group ($N=13$) had a mean use score of 1·54 as compared with 3·26 for the experimental group ($N=8$). $S\sqrt{(\text{var } S)}=2{\cdot}2$, i.e. significant at 5% level. Their mean recall score $\times 2$ was 1·76. By comparing these figures with those in Table 1, it can be seen that although the older subjects were relatively worse at recalling the propaganda, their mean use score was relatively high. Although they did not remember the posters there is ample evidence that they had learned the information contained in them.

(3) *The effect of time on the use and recall of propaganda—the length of time between display of the propaganda and measurement of its effectiveness*

Hitherto, the photograph test had been presented immediately after the display of propaganda. This test was now given at intervals of from 1 to 14 days after the display of Pictorial Propaganda. Subjects visited the laboratory twice. On the first occasion they saw the Pictorial Propaganda in the waiting room and were asked to interpret a few non-test photos, and on the second occasion they were given the usual photograph test and were asked to describe the posters that they had seen on their previous visit. Sixty different student subjects were tested (thirty men and thirty women).

The mean use and recall scores of these subjects have been broken down into drivers and non-drivers and are shown in Table 2. It will be seen that there was again a contrast between the use scores of drivers and non-drivers. (On the test photos there was no significant difference between the mean scores of drivers/non-drivers in a control group of students who had not been subject to propaganda.) The drivers applied the information from the propaganda more than the non-drivers. On the other hand, the non-drivers tended to recall slightly more than the drivers, although there was no very marked or uniform difference between the recall scores of the two groups. It was decided that in any further comparisons the scores should be weighted for driving/non-driving (on the assumption that each group would contain 50% drivers and 50% non-drivers.) The weighted means are also shown in Table 2, from which it can be seen that there is a steady decline with time in the performance of the subjects both on the use and recall scores. The small rise in recall scores after 7 days can be accounted for entirely by the scores of the five women in the group, who, it is suspected, had discussed the posters. The mean score of the men after 7 days was 0·34. It is interesting to note, however, that whatever the reason for this rise, the use scores did not rise correspondingly.

When significance tests were applied to the data (unweighted) it was found that the propaganda no longer had any significant effect after 1 day. It can be seen from the table that while there was still some use of the information 2 weeks after the propaganda display, there was practically no ability to remember the posters. The relationship between use and recall scores can be seen most clearly from a consideration of the effects of individual posters. For example, while subjects were not able to remember poster no. 4 at all after 2 weeks, the group was still better than the control group after this period at the photo test. There was a widening disparity with time between recall and use (Text-fig. 1).

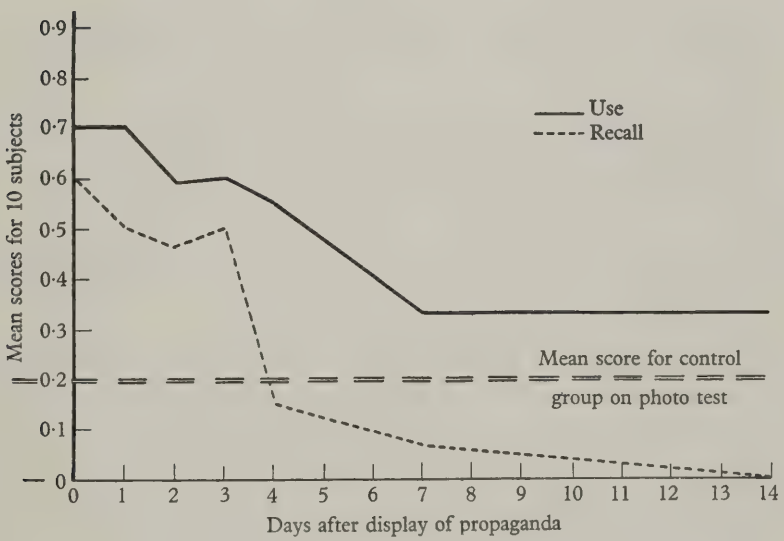
Let us now consider two general points in relation to all the preceding data. First, how much of the propaganda was recalled without being used? And secondly, how much of the propaganda was used but not recalled during the test?

Table 2. *The effect of time on the use and recall scores*

Students (five male and five female in each group)

Time interval between display of propaganda and test (days)	Mean use scores		Weighted mean use scores
	Drivers	Non-drivers	
1	4·25	2·0	3·13
2	3·0	2·6	2·8
3	2·6	2·5	2·55
4	2·7	2·3	2·5
7	2·0	2·1	2·05
14	2·3	2·0	2·15
	Mean score of Control group		1·8

Time interval between display of propaganda and test (days)	Mean recall scores		Weighted mean recall scores
	Drivers	Non-drivers	
1	2·0	1·4	1·7
2	1·0	1·4	1·2
3	0·6	1·2	0·9
4	0·3	0·6	0·45
7	0·7	1·1	0·9
14	0·3	0	0·15



Text-fig. 1. Mean use and recall scores (weighted for car driving) for one poster and photo (no. 4).

(4) *Propaganda recalled but not used*

An analysis of the responses of individual subjects shows that a number of posters were recalled although the relevant photos were not responded to accurately in the test. Table 3 shows the amount recalled but not used by both naval ratings and students.

Table 3. *Amounts recalled but not used by naval ratings and students*

	Percentage recalled of possible total responses	Percentage recalled but-not-used responses of possible total responses
Naval ratings:		
Pictorial	43	10
Horror	43	15
Verbal positive	32	5
Verbal negative	48	15
Students:		
Pictorial	52	11
Horror	37	3
Verbal positive	47	13
Verbal negative	32	8

(5) *Propaganda used but not recalled*

The number of posters which was used without being recalled is not subject to direct measurement. Although a number of photos was responded to accurately without the posters being recalled, these may well have been ones which subjects would have had correct in any case, i.e. without propaganda (cf. the control groups).

An attempt has been made, however, to determine very approximately how much of the propaganda was used without being recalled. In the photo test it was possible for a subject to make any of the following combinations of responses:

Photo correct + relevant poster recalled correctly = y responses,
Photo correct + relevant poster recalled incorrectly,
Photo correct + relevant poster not recalled = z responses,
Photo incorrect + relevant poster recalled correctly,
Photo incorrect + relevant poster recalled incorrectly,
Photo incorrect + relevant poster not recalled.

If the number of subjects in the group = x, and the number of photographs given to each subject = a, then the fraction $100\,\Sigma z/ax - \Sigma y$ can reasonably be compared with the percentage of responses in the control groups, i.e. we can ask the question 'other than those known to have been influenced by the posters, were there more in the experimental groups who Used the information than might have been expected from inspection of the control groups?' Table 4 shows a number of these comparisons. It can be seen that: (i) The value of $100\,\Sigma z/ax - \Sigma y$ for the student groups is remarkably consistent—whether it is for different types of propaganda or for the Pictorial Propaganda after different time intervals. (ii) The value of $100\,\Sigma z/ax - \Sigma y$ for the student groups differs very little from the percentage correct responses in the control group. (iii) In contrast to the student group, there was considerable variation between the naval rating groups. (The reason for the low score on the Horror group was that a number of posters were inaccurately recalled.) (iv) The over-sixties group, whilst having a control score comparable with that of the naval ratings and students, had a value for $100\,\Sigma z/ax - \Sigma y$ higher than for any other group, indicating that this group Used more posters without Recalling them than did any other group.

There is some evidence that with the student groups all types of propaganda were likely to be recalled as well as used. With the naval-rating groups there was a greater tendency to use information which could not be recalled, especially with Pictorial Propaganda. The over-sixties group showed the greatest tendency towards using information which could not be recalled.

Table 4. *Values of* $100\,\Sigma z/ax - \Sigma y$ *for different subjects and for different types of propaganda*

Type of propaganda	Naval ratings	Students	Over sixties
Pictorial (immediate)	40	30	44
Horror	12	32	—
Verbal positive	29	33	—
Verbal negative	30	34	—
Pictorial:			
1 day	—	31	—
2 days	—	28	—
3 days	—	30	—
4 days	—	36	—
7 days	—	32	—
14 days	—	32	—
Control group percentages of correct responses to the total possible responses	24	30	26

(6) *Attention to the propaganda and introspective evidence*

Finally, we must consider the introspection of the subjects to see whether any light is thrown on the problem of why some subjects used (or recalled) more information than they recalled (or used).

It was evident that even the most successful of the various types of propaganda was only used about half as much as it was possible to use it. One possible explanation is that the subjects did not look at the propaganda. Only two subjects of over 200 who were tested, however, said that they had not seen any posters. Their scores were omitted from the results. Proof that the posters were 'looked at' is that every subject described them as a homogeneous group. Many who could not recall the implication of the posters were able to say that 'two were for cyclists, one for pedestrians, etc'.

There is adequate evidence, however, that although all the subjects looked at the propaganda, some did not in fact perceive it. Many of the picture posters were described wrongly, e.g. 'Face oncoming traffic' was often interpreted as 'Don't cross in front of a bus'. Many subjects said that they looked at the posters merely from an artistic or critical point of view. They seemed to have perceived structural items, but not the total concept of the picture. Some had consciously rejected any idea of reading the slogans and thus of being able to understand the implication of the posters, e.g. 'I thought "Road Safety again" and didn't look any further.'

On the other hand, there were subjects who read and understood the implications of the posters, but who failed to change their attitude towards the particular problem, e.g. 'I read them all, but they didn't have the slightest effect on the Photo test, as I knew all the points already' (use score 1/6). There were still a number of subjects, however, who read, recalled and expressed appreciation of the propaganda and yet who failed to be

influenced by it at all in the photo test, e.g. 'Propaganda does a lot of good—the more the better' (use score 1/6). It should be noted that none of the subjects was aware of the object of the experiment.

IV. Discussion

An analysis of the data in this experiment shows that, although the type of propaganda had little effect on the results, the relationship between recall and use of the propaganda varied in amount and direction both within and between different groups of subjects. We cannot account for these results by postulating that either recall or use was more sensitive as a measure of effectiveness. For although some subjects tended to use more propaganda than they could recall (e.g. the car drivers and over-sixties), others used most of the propaganda which they were able to recall (e.g. non-drivers). Also, there were some subjects who tended to Recall more than they used. One argument which might be put forward is that individual items of the propaganda were recalled, although the propaganda was not understood as a meaningful whole and, in consequence, did not affect behaviour. As Vernon (1951) states '. . . memorization of particular words and phrases, and even of whole statements, can be achieved without any assimilation of the general argument or understanding of the ultimate meaning of a written passage'. It is hardly likely, however, that the one sentence slogans used in this experiment could have been recalled without being understood. A second argument is that some of the propaganda might have been understood and recalled but not accepted as useful information. This is unlikely to have accounted for the performance of many subjects, as so many of them voluntarily expressed approval of the slogans which they had recalled but not used.

Welborn & English (1937) have summarized a considerable amount of experimental work which has been designed to show the difference or similarity between 'verbatim' and 'substance' learning. Most of the tests of substance learning, however, have either required a verbal or written précis of the original or have been direct question objective tests. The influence of the learning on subsequent behaviour has not been measured and the results cannot therefore be adequately compared with our own. Katona (1940) developed the distinction between learning by memorizing and learning substance one stage further in the field of problem solving. But in none of this previous work has verbal material been learned and then the information transferred to a non-verbal task. Szekely (1950), however, has examined the question of whether a functional connection exists between insightful previous learning and subsequent problem solving by studying the effects of two different ways of learning verbal material. While one group (the 'modern' method group) learned the material by incorporating it into their activity, the others ('traditional' method group) merely learned the substance. The learning material was the solution of a simple mechanical task. The solution of the final task involved the knowledge of the principles which had been learned in the first task. He found that the frequency of solution in the final test was considerably higher among the subjects who learned the concepts according to the so-called modern learning method. These subjects, however, were seldom able to verbalize the learning material. Subjects who learned by the traditional method verbalized more frequently than subjects who had learned by the modern method. His results are thus very much akin to our own.

In the work of Katona and Szekely it was clear why some subjects responded in one way to the learning material and other subjects responded in quite a different way. It was

because the experimenter imposed a manner of learning upon the subjects. In our experiments, however, subjects were not asked to learn the material at all. It is possible, however, that different 'sets' towards learning existed between our different groups which determined *how* they learnt and what they extracted from the material. These different methods of learning may in turn have determined (as in Szekely's experiments) the final form of response. Let us examine whether individual differences existed that might have accounted for different 'sets' towards learning.

The greatest significant difference in our results was between drivers and non-drivers, the drivers using far more of the information in relation to the amount they recalled than non-drivers. Car-drivers have, in learning to drive, assimilated road safety information in relation to its being incorporated into their car-driving behaviour. These learning periods have in fact been very similar to the activity learning or modern method used by Szekely in his experiment—information was organized and learnt in relation to its being used. It is likely, therefore, that car drivers in repeatedly organizing material in this way during their experience of learning-in-order-to-use and of using are building up a way of apprehending and reacting to information in a particular situation. Thus when the car-drivers are presented with further learning material such as this incidental propaganda the response seems likely to be organized consistently with previous responses. They will in fact have a set towards learning-in-order-to-use—but not necessarily to recall. As Postman & Senders (1946) have suggested, sets not only determine the course of learning but may themselves be the result of previous learning.

Our hypothesis now needs to be considered in the light of the performance of our other groups. We noted that students tended to recall more information in relation to the amount used than the naval ratings. It was also apparent that the older subjects had very poor recall scores, although their use scores were relatively high. (This latter finding is in contrast to the assumption which has generally been made, and which has been based on the results of memory tests, that older subjects learn less well than younger ones.)

Where there has been no previous learning specific to the new learning situation (in this case, road safety regulations), it seems likely that individuals will respond in the way they have been most accustomed to respond in other learning situations. As a product of our educational system, learning information (Szekely's traditional method) and memorizing are likely to be predominant forms of responses. On these assumptions, it is hardly surprising that the student subjects should have a high recall score. They have continuously exercised the 'traditional method' of learning and response. The formal education of the naval ratings, however, ceased at 14 or 15 years of age, since when 'learning by activity' is likely to have been predominant in their training in the navy. Again we can postulate that the older people through their experience over many years have developed preferred persistent responses in a learning situation. This particular group of over-sixties had a low educational attainment and were mostly artisans. Their learning had obviously been orientated more towards 'activity'.

Our hypothesis (which will of course have to be tested in relation to other learning situations before its generality can be assumed) is that the manner in which incidental information is retained (either in memory or for use in relevant action) will be determined at the time of learning and conditioned by previous learned responses. If this is so, then it seems that for propaganda to have the optimum effect on behaviour the individual should previously have organized appropriate behaviour into his learning experience.

Propaganda should in fact be used as a reinforcement of some more active learning. The results also yield implications for educational practice. In the past both teachers and psychologists have focused their attention on the importance of learning by repetition or learning substance. It is clear that these are not always the most effective ways of learning if the application of such learning is the criterion of effectiveness. What seems of even greater importance is that these methods of learning may well become habitual responses and so determine the course and manner of future learning. The teacher who assumes that ability to reproduce learned material is indicative of the ability to apply the information usefully is apparently making a false assumption. What is worse is that he seems to be prejudicing the individual's chances in future learning situations.

References

Belbin, E. (1953). The effect of propaganda on recall, recognition and behaviour. Cambridge Univ. Library, unpublished Ph.D. thesis.

Belbin, E. (1956). The effects of propaganda on recall, recognition and behaviour. Part I. The relationship between the different measures of propaganda effectiveness. *Brit. J. Psychol.* **47**, 163–74.

Katona, G. (1940). *Organising and Memorising.* New York: Columbia University Press.

Postman, L. & Senders, U. L. (1946). Incidental learning and generality of set. *J. Exp. Psychol.* **36**, 153–65.

Szekely, L. (1950). Productive processes in learning and thinking. *Acta psychol., Hague*, **7**, 388–407.

Vernon, M. D. (1951). Learning and understanding. *Quart. J. Exp. Psychol.* **3**, 19–23.

Welborn, E. L. & English, H. L. (1937). Logical learning and retention. *Psychol. Bull.* **34**, 1–20.

Whitfield, J. W. (1947). Rank correlation between two variables, one of which is ranked, the other dichotomous. *Biometrika*, **34**, 292–296.

(*Manuscript received* 6 *June* 1955)

DISPLAY-CONTROL RELATIONSHIPS ON CIRCULAR AND LINEAR SCALES

By N. E. LOVELESS

The Nuffield Department of Industrial Health, University of Durham
(King's College, Newcastle upon Tyne)

The variation in tracking performance between different quadrants of a circular scale has been shown to be attributable mainly to the ambiguity of display-control relationships on this scale, in that the motion of the pointer may be viewed as either rotary or translatory. The dominant tendency is to view it as rotary, a clockwise control movement being associated with a clockwise pointer movement; but there is a significant secondary tendency to view it as translatory, a clockwise control movement being associated with a displacement of the pointer upwards or to the right, according to the quadrant in which the target lies. The two tendencies conflict when the target is in the lower and right-hand quadrants, where performance is consequently poorer. Performance is also affected by position in the visual field as such.

It has been verified that a clockwise control movement is expected to send the pointer to the right on a horizontal linear scale and upward on a vertical linear scale. Performance on the horizontal scale is no better than that obtained on the circular scale under optimal conditions, but the unambiguity of the linear scale renders it preferable where erroneous reactions are to be avoided.

I. Introduction

Workers in the field of equipment design early recognized that the problems of display and control are not independent and, in particular, that, in the frequent case in which the display involves a moving element (such as a pointer), the directional relationship between pointer movement and control movement may become a variable which has a considerable influence on the speed and accuracy of operation. Where, for example, a pointer and a control lever move linearly in the same direction, the operator appears to expect that a given control movement will result in a pointer movement of the same sense ('direct' relationship), and may be disturbed if it results in one of opposite sense ('reversed' relationship).

It has been found that, in many cases, there is such a 'preferred' directional relationship between control and display. Where the task is not too demanding, the preferred relationship may yield no better performance than its opposite (Grether, 1947); but where the task is more stressful, as when it demands speed or division of attention, the directional relationship may become important (Mitchell & Vince, 1951). Such experimental evidence as is available (Mitchell & Vince, 1951) indicates that, with sufficient practice, performance with a non-preferred arrangement may be brought to as good a level as that obtained with a preferred relationship. The lengthier training period may, of course, be something of a disadvantage; but more serious, perhaps, is the fact that although training may succeed in reversing an existing habit, the operator is liable to revert to his earlier reaction under conditions of stress (Simon, 1951). It is therefore the more necessary, where stress is anticipated and the result of a mistake may be serious, to design equipment so that direction-of-motion relationships conform to the existing expectations of the majority of operators.

Some situations are found to be so unfamiliar that no stereotype exists, as when display

and control move in different planes (Carter & Murray, 1947; Gardner, 1950; Mitchell & Vince, 1951; Warrick, 1947). These situations are ambiguous in the sense that an untrained operator is equally likely to make either of two opposing reactions. Other situations, such as the circular scale with which we are here concerned, are ambiguous in another sense: namely, that they may be susceptible of more than one interpretation.

II. Display-control relationships on the circular scale

Despite the wide use of circular scales, rather little attention has been paid to the display-control relationships associated with them.

Fitzwater (1948) studied a compensatory pursuit task in which the operator used a control lever to maintain the pointer on the '12 o'clock' mark of a circular scale, from which it was periodically displaced by externally produced disturbances. He found that a significantly better score was obtained when a clockwise movement of the lever resulted in a clockwise rotation of the pointer than when it caused the pointer to rotate counter-clockwise.

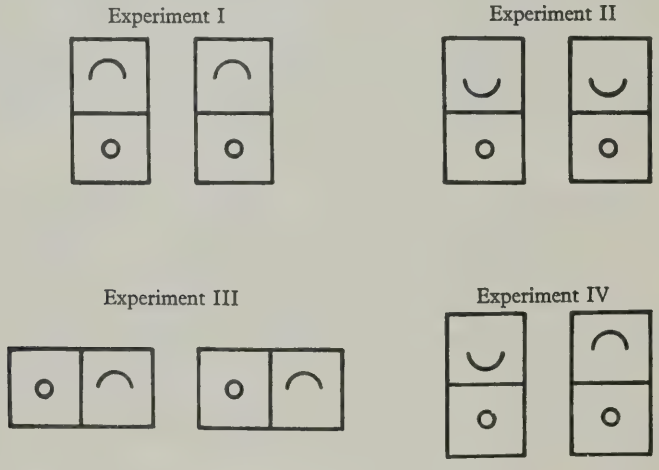


Fig. 1. Arrangements of the control and indicator units used in Warrick's experiment (1949).

Warrick (1948) used a printed test, asking subjects in which direction they would turn a circular control knob to correct a given deviation on a quadrantal scale. The upper, lower, left and right quadrants were studied, each in combination with a control which might be above, below, to right or left of it, thus giving sixteen arrangements in all. The clockwise/clockwise hypothesis was supported at the 1% level of confidence, differences between the arrangements being insignificant.

In a further study by Warrick (1949), subjects alternately adjusted two semicircular indicators by means of rotary control knobs. The relative position of indicator and control was varied in four experiments, as shown in Fig. 1. Within each experiment, the control-display relationship was varied, direct-direct, reversed-reversed, direct-reversed, and reversed-direct combinations (*DD*, *RR*, *DR* and *RD*) being presented in balanced order.

The results from the first three experiments showed that when two indicator-control units with like arrangements are operated alternately, operators perform more rapidly

and consistently if the motion relationships of the units are the same—that is, the *DD* and *RR* combinations are superior to the *DR* and *RD*. Moreover, a difference was found between the consistent combinations, direct drive being superior to reversed (*DD* to *RR*). The clockwise/clockwise hypothesis is thus further supported.

The fourth experiment, in which the panel arrangements were unlike, yielded rather different results. *DD* was no better than *DR* and *RD*, while *RR* was much worse. Warrick's interpretation is that when panel arrangements are unlike, consistent motion-relationships do not help, and none of the relationships are as effective as when panel arrangements are alike.

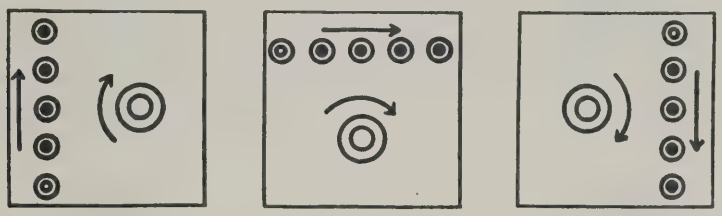


Fig. 2. Response preferences determined by Warrick (1947).

Murrell (1952) comments that Warrick's interpretation of his findings appears to be influenced by a general principle of control-display relationships which he enunciated as the conclusion of a previous experiment using linear scales (Warrick, 1947), namely, that a pointer is expected to move in the same direction as that part of the control nearest to the display (see Fig. 2). This principle, when carried over to the most frequently used upper segment of a circular dial, reinforces the common-sense assumption that the expected relationship between a knob and a circular dial is clockwise/clockwise (Fig. 3*a*). But Murrell suggests that in the bottom segment of the scale, the expected relationship will not necessarily be clockwise/clockwise; this would, in fact, be contrary to the first conclusion above, since the pointer would then move in the opposite sense to the nearest part of the control (Fig. 3*b*). He points out that in Warrick's Exp. IV, the best arrangement was in fact *RD*, in which a right-left pointer movement in the bottom half of the dial corresponds with a right-left (anticlockwise) movement of the upper part of the control.

Murrell elsewhere (1951) puts forward his own principle of display-control relationships. Research, he says 'has established beyond reasonable doubt that when a circular dial and control are in the same plane the operator expects the pointer to move as if it were pivoted at the centre of the control'. The point of this formulation is evidently to include a reversal in the bottom segment of the circular scale.

Such a reversal is not, however, established 'beyond reasonable doubt'. The difference in Warrick's Exp. IV is not significant, and cannot in any case be considered uninfluenced by the alternation of tasks. It is, however, true that Warrick ignores the possible significance of unequal performance in different segments of the scale, even though there is some suggestion of this in his own Exps. I and II. Evidence of such irregularities would require modification of the clockwise/clockwise hypothesis.

Graham (1952), using a compensatory pursuit task with a control knob below the dial,

found differences in performance with choice of target. Speed and accuracy were better when the target was at the top of the scale than when it was at the bottom, right or left cardinal points. She suggested that this might be due to an ambiguity in the control-display relationships.

Similarly, Fitts & Simon (1952), in a dual pursuit task, found performance to be superior in the top and left-hand quadrants of the dials. They point out that similar differences have been found in other types of investigation. Dallenbach (1923) found that

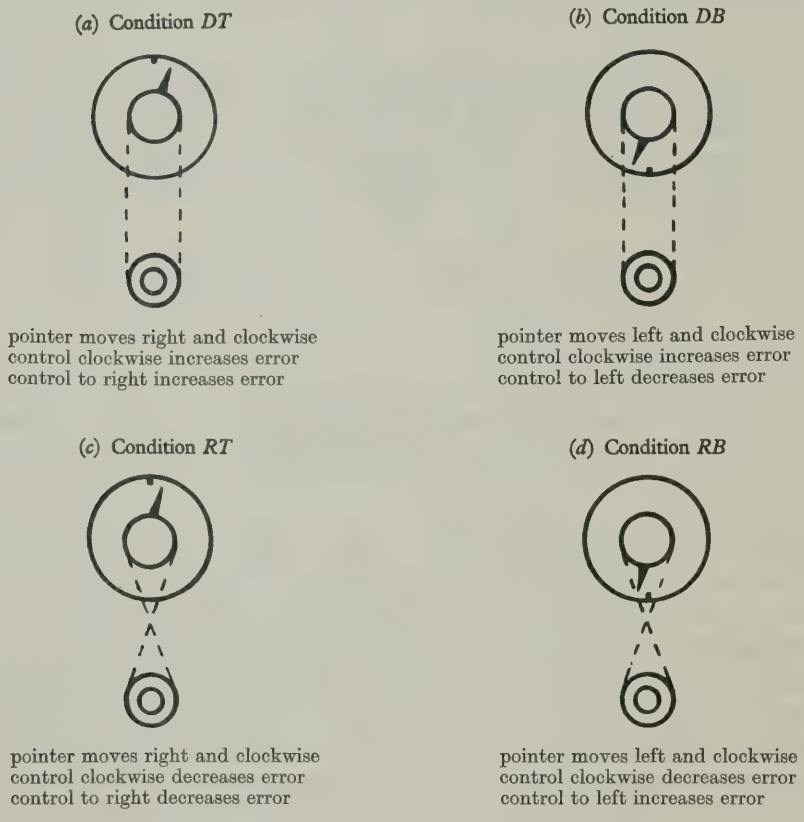

(a) Condition DT
pointer moves right and clockwise
control clockwise increases error
control to right increases error
(b) Condition DB
pointer moves left and clockwise
control clockwise increases error
control to left decreases error
(c) Condition RT
pointer moves right and clockwise
control clockwise decreases error
control to right decreases error
(d) Condition RB
pointer moves left and clockwise
control clockwise decreases error
control to left increases error

Fig. 3. Display-control relationship on the circular scale.

lights appeared brighter in these quadrants than when below or to the right of the fixation point. Eye fixations have been found to be most frequent in this area. Connell & Grether (1948) and White (1949) have shown that dials are read more quickly and accurately in these sectors.

It is not clear how far, if at all, Fitts & Simon consider that such a factor of position in the visual field might account for their own results. Certainly they put forward an alternative explanation, that 'direction-of-motion cues are ambiguous when rotary motion occurs in these (bottom and right) sectors. In order to account for this ambiguity, we assume that a moving stimulus object, such as an instrument pointer, is responded to in accordance with either a "rectangular-coordinate hypothesis" or a "polar-coordinate

hypothesis". If we assume that the two hypotheses lead to the same motor or verbal response for a specified angular position of the pointer, then the two hypotheses will lead to opposite responses when the pointer has passed through an 180° rotation (i.e. on the reverse side of the dial). As an illustration, a clockwise rotation at 9 o'clock is "up" and at 3 o'clock is "down", according to the rectangular hypothesis'.

Fitts & Simon are careful to stress that the ambiguity of a display-control relationship may be affected by the type of control used and its placement—for example, when a lever is used instead of a rotary control, its orientation will be important (Simon, 1952). Again, the ambiguity effect may well be enhanced by the use in their experiment of two spatially separated pointers. But they feel that, since only operators responding in accordance with the rectangular hypothesis would show a poorer level of performance in the 3 o'clock and 6 o'clock positions, 'this provides indirect evidence that most of our subjects must have responded at times on the basis of this hypothesis, even though the displays and controls provide a rotary or polar-coordinate system'.

III. Analysis of performance on a circular scale

It is suggested, then, that in the common case in which the control is a knob in the same plane as the display, the display-control relationship may be ambiguous, in that the operator may regard the pointer movement either as rotary, or, over a limited segment of the scale, as translatory. If he acts on the first hypothesis, then he may, for example, expect a clockwise control movement to move the pointer clockwise; whereas, if he acts on the second assumption, he may expect it to move to the right. When the target is at the top of the scale, both expectations lead to the same response; but, at the bottom of the scale, they lead to antagonistic response tendencies, conflict between which will be reflected in poorer performance.

It is clear that those differences in performance in different sectors of the circular scale which have so far been reported may be accounted for along these lines. Unfortunately, they may equally well result simply from the different locations of the target within the visual field, an hypothesis which finds some support in the observations quoted by Fitts & Simon. The experimental evidence to date does not allow us to decide between these two explanations. Such a decision requires an examination of performance on diametrically opposite targets and with both directions of drive. This may be seen from the following analysis.

Suppose that tracking error scores are obtained under four conditions: with targets at the top and bottom of the scale, in each case with both direct and reversed drive. Let the four scores be denoted by *DT*, *RT*, *DB*, *RB*.

It may be supposed that these scores will be affected by the two direction-of-motion expectations suggested above. Let the scores be presumed to contain components 'r' and 't' corresponding to the 'rotary' and 'translatory' reaction tendencies. The scores may also be affected by choice of target, in that performance at the top of the scale may be inherently more easy or difficult than at the bottom. Let there be a component 'p' corresponding to this position factor. Finally, the scores will be affected by a number of constant factors whose effects will be included in the general overall mean 'm'.

Any particular score may then be regarded as the resultant of four components:

m the general overall mean.

p the position term, arising from differences between targets as such, and defined as positive if the top target produces the greater error.

r the term arising from the 'rotary tendency', defined as positive when a clockwise displacement of the pointer evokes a clockwise correction.

t the term arising from the 'translatory tendency', defined as positive when a displacement of the pointer to the right evokes a clockwise correction.

Now, under the condition DT (Fig. 3*a*), the components 'r' and 't' will operate in the same direction. For example, the pointer moves to the right (clockwise). If 'r' is positive, the operator will tend to move the control clockwise, so increasing his error. Similarly, if 't' is positive, he will again tend to move the control clockwise, increasing his error score. Hence the score under condition DT is given by

$$DT = m + p + r + t.$$

Conversely, under the condition RT (Fig. 3*c*), positive 'r' and 't' components both operate to decrease the scores, whence

$$RT = m + p - r - t.$$

On the bottom target, however, 'r' and 't' become antagonistic. When direct drive is used (Fig. 3*b*), positive 'r' increases the error and positive 't' decreases it, so that

$$DB = m - p + r - t.$$

And when the drive is reversed (Fig. 3*d*),

$$RB = m - p - r + t.$$

Solution of these equations gives

$$\begin{aligned} m &= \tfrac{1}{4}(DT + RT + DB + RB), \\ p &= \tfrac{1}{4}(DT + RT - DB - RB), \\ r &= \tfrac{1}{4}(DT - RT + DB - RB), \\ t &= \tfrac{1}{4}(DT - RT - DB + RB). \end{aligned}$$

Thus, from such a set of scores, we may estimate the strength of these components. The statistical significance of the last three factors may be estimated from the 'between targets', 'between drives', and 'drives × targets interaction' terms of an analysis of variance of the scores.

Experiments which allow of such analysis should indicate what factors are needed to account for observed variations of performance in different sectors of a circular scale.

IV. Experiment one: upper and lower quadrants

This experiment was designed to permit the analysis described above; that is, a comparison was made of performance on targets at the top and bottom cardinal points of a circular scale, with both directions of drive.

The task used was compensatory tracking. The display consisted of a scale 2·75 in. in diameter, with 10 × 10 white-on-black graduations. Vertically below it was a control knob 2·5 in. in diameter, connected through a differential gear to the pointer. The second input to the differential gear was obtained from a cam follower, which produced step-type

disturbances at an average rate of one every 2·1 sec. The subject's task was to counteract these disturbances so as to maintain the pointer on a given target. A voltage proportional to the deviation of the pointer from the target was obtained from a potentiometer, and integrated against time by an electronic device. The score so obtained was the 'tracking error', measured in arbitrary units. The scores were subsequently normalized by a logarithmic transformation.

The subjects were tested individually. Before the test, the subject was allowed to handle the control knob for 15 sec. to discover the display-control relationship. This was followed by 60 sec. practice. The test run lasted 90 sec., but the first and last 15 sec. were not scored.

The subjects were seventy-two aircraftmen of the Royal Air Force, their ages ranging from 18 to 26 years. All had visual acuity of 20/20 or better as tested by the Snellen chart. They were divided into groups of eighteen subjects, each group being tested under one of the four experimental conditions.

The mean scores obtained under the four conditions, estimates of the factors presumed to account for these scores, and of the significance of these factors as estimated by an analysis of variance, are shown in Table 1. Tests of the significance of differences between means show that the error score obtained under condition *DT* is smaller than that obtained under the other conditions; no other differences are significant at the 5 % level of confidence. All the terms of the analysis of variance are significant at this level.

Table 1. *Mean scores, score components, and analysis of variance—Exp. I*

(*a*) Mean Scores

	Means	Differences			
		DT	*DB*	*RT*	*RB*
DT	154·1	—	14·8*	20·8*	21·2*
DB	168·9	—	—	6·0	6·4
RT	174·9	—	—	—	0·4
RB	175·3	—	—	—	—

* Significance at 5 % c.l.

(*b*) Score components

$m = 168{\cdot}30$
$p = -3{\cdot}80$
$r = -6{\cdot}80$
$t = -3{\cdot}60$

(*c*) Analysis of variance

Source of variation	Sums of squares	Degrees of freedom	Mean squares	*F*	Significance (5 % level)
Between drives	3335	1	3335	17·28	*
Between targets	1043	1	1043	5·40	*
Drives × targets	953	1	953	4·98	*
Residual	13138	68	193	—	—
Total	18469	71	—	—	—

It will be seen that although both direction-of-motion factors are present, '*r*' appears to be the dominant tendency—that is, subjects react primarily to angular motion of the pointer. There is, however, a significant secondary tendency to respond in terms of 'left' and 'right', a clockwise control movement being expected to move the pointer to the right. Finally, there is a tendency for performance at the top of the scale to be inherently superior to that at the bottom.

V. Experiment two: left-hand and right-hand quadrants

This experiment involved a comparison of performance on targets at the left-hand and right-hand cardinal points of the scale, again with both directions of drive, so that the four experimental conditions were *DR, DL, RR, RL.* In all other respects the design of the experiment was identical with that first described. The results are presented in Table 2.

The variance of *DR* is significantly smaller than that for the other three conditions, and the differences between means have therefore been tested by Behren's method (Fisher & Yates, 1948). The difference between *RL* and *RR* is not significant at the 5% level of confidence. *DR* is superior to either of the reversed drive conditions, and *DL* to all other conditions. The analysis of variance shows that the 'between drives' and 'drives × targets interaction' terms are significant at the 5% level.

Table 2. *Mean scores, score components, and analysis of variance—Exp. II*

(*a*) Mean scores

	Means	Differences			
		DL	*DR*	*RL*	*RR*
DL	154·7	—	8·0*	29·7*	24·3*
DR	162·7	—	—	21·7*	16·3*
RL	184·4	—	—	—	5·4
RR	179·0	—	—	—	—

* Significance at 5% c.l.

(*b*) Score components

$m = 170{\cdot}20$
$p = -0{\cdot}65$
$r = -11{\cdot}50$
$t = -3{\cdot}35$

(*c*) Analysis of variance

Source of variation	Sums of squares	Degrees of freedom	Mean squares	*F*	Significance (5% level)
Between drives	9499	1	9499	47·50	*
Between targets	31	1	31	0·15	—
Drives × targets	807	1	807	4·08	*
Residual	13447	68	198	—	—
Total	23784	71	—	—	—

Both direction-of-motion factors are thus present, and the 'rotary tendency' again appears to be dominant, a clockwise control movement being expected to move the pointer clockwise. But again there is a significant secondary tendency, this time to respond in terms of 'up' and 'down', a clockwise control movement being expected to move the pointer upwards.

It will be seen that there is no evidence in this experiment for an inherent superiority of the left-hand target. Differences in scores are attributable solely to the direction-of-motion factors.

VI. Discussion on the circular scale

It appears then that differences in performance on the circular scale can be accounted for by three factors:

(*a*) A general tendency for performance to improve as the target is moved from the

bottom to the top of the scale. The reason for this is uncertain. It is possible that the position of the target relative to the subject's eye-level is involved; but if so, this variable is more important than has generally been supposed. The evidence cited above predicted an improvement from right to left also; this did not appear in the present results.

(*b*) An expectation that a clockwise control movement will result in a clockwise pointer movement. This finding is supported by other evidence.

(*c*) An expectation that a clockwise control movement will move the pointer to the right or upwards, according to the target used. The movement to the right conforms to Warrick's principle that a pointer is expected to move in the same direction as that part of the control nearest the display. It also agrees with what is known of direction-of-motion relationships on the horizontal linear scale. In respect of the tendency to expect the horizontally aligned pointer to move upwards, it would seem desirable to seek confirmation in the case of a vertical linear scale with control knob beneath it; no evidence bearing on this case has been found in the literature.

VII. Experiment three: linear scales

This third experiment was therefore intended primarily to determine the optimal direction-of-motion relationship on a vertical linear scale. The optimal relationship on a horizontal linear scale was not in doubt, but it was decided to include such a scale also in order to secure comparative measures of performance. There were thus four experimental conditions, according as a clockwise control movement displaced the pointer upward, downward, to right or to left (conditions *U*, *D*, *R*, *L*). The length of the scale was equivalent to that of the circumference of the circular scale; the target was at its central point. In other respects the experimental design was the same as in the first two experiments. The results are shown in Table 3.

Table 3. *Mean scores, score components and analysis of variance—Exp. III*

(*a*) Mean scores

	Means	Differences			
		D	*U*	*L*	*R*
D	174·1	—	11·1*	4·2	20·1*
U	163·0	—	—	15·3*	9·0*
L	178·3	—	—	—	24·3*
R	154·0	—	—	—	—

* Significance at 5% c.l.

(*b*) Score components

Vertical scale: $m = 168{\cdot}6$, $t = -5{\cdot}6$

Horizontal scale: $m = 166{\cdot}2$, $t = -12{\cdot}2$

(*c*) Analysis of variance

Source of variation	Sums of squares	Degrees of freedom	Mean squares	*F*	Significance (5% level)
Between scales	103	1	103	0·66	—
Between drives	5653	1	5653	36·47	*
Drives × scales	800	1	800	5·16	*
Residual	10514	68	155	—	—
Total	17070	71	—	—	—

The scores for the different conditions differ significantly in variance, and allowance has been made for this in testing the differences between means. The difference between *D* and *L* is not significant at the 5% level of confidence. *U* is superior to both these conditions, and *R* to all other conditions. This verifies the predicted association of a clockwise control movement with an upward pointer movement on the vertical scale and a movement to the right on the horizontal scale; performance under the latter condition is superior to that under the former. The analysis of variance shows that this is due to a stronger association of display and control movements on the horizontal scale rather than to any difference in mean performance on the two scales.

VIII. General Discussion

Measures of speed and accuracy of tracking have now been obtained under both display-control relationships for horizontal and vertical scales and for targets at the four cardinal points of the circular scale. The results are brought together in Table 4, where the different arrangements are arranged in order of magnitude of the means. The variances are just significantly heterogeneous at the 5% level of confidence, and appropriate allowance has been made for this in testing the significance of differences between means.

Table 4. *All scales*: *means, variances and differences between means*

				Differences between means										
Scale	Condition	Mean	Variance	*DT*	*DL*	*DR*	*U*	*DB*	*D*	*RT*	*RB*	*L*	*RR*	*RL*
Horizontal	*R*	154·0	57·6	0·1	0·7	8·7*	9·0*	14·9*	20·1*	20·9*	21·3*	24·3*	25·0*	30·4*
Circular	*DT*	154·1	129·0	—	0·6	8·6*	8·9*	14·8*	20·0*	20·8*	21·2*	24·2*	24·9*	30·3*
Circular	*DL*	154·7	152·9	—	—	8·0*	8·3*	14·2*	19·4*	20·2*	20·6*	23·6*	24·3*	29·7*
Circular	*DR*	162·7	65·0	—	—	—	0·3	6·2	11·4*	12·2*	12·6*	15·6*	16·3*	21·7*
Vertical	*U*	163·0	79·4	—	—	—	—	5·9	11·1*	11·9*	12·3*	15·3*	16·0*	21·4*
Circular	*DB*	168·9	139·0	—	—	—	—	—	5·2	6·0	6·4	9·4*	10·1*	15·5*
Vertical	*D*	174·1	330·6	—	—	—	—	—	—	0·8	1·2	4·2	4·9	10·3
Circular	*RT*	174·9	274·2	—	—	—	—	—	—	—	0·4	3·4	4·1	9·5
Circular	*RB*	175·3	230·7	—	—	—	—	—	—	—	—	3·0	3·7	9·1
Horizontal	*L*	178·3	150·8	—	—	—	—	—	—	—	—	—	0·7	6·1
Circular	*RR*	179·0	266·3	—	—	—	—	—	—	—	—	—	—	5·4
Circular	*RL*	184·4	291·1	—	—	—	—	—	—	—	—	—	—	—

* Significance at 5% c.l.

The results of this test suggest that the twelve arrangements investigated may be classified into three groups in the following order of decreasing merit:

(A) Horizontal linear scale, clockwise control movement sending pointer to right.

Circular scale, target at top, clockwise control movement sending pointer clockwise and to right.

Circular scale, target at left, clockwise control movement sending pointer clockwise and upwards.

(B) Vertical linear scale, clockwise control movement sending pointer upwards.

Circular scale, target at bottom, clockwise control movement sending pointer clockwise and to left.

Circular scale, target at right, clockwise control movement sending pointer clockwise and downwards.

(C) Arrangements with control display relationships the opposite of above.

We have accounted for the way in which performance on the circular scale varies with the position of the target by stressing the ambiguity of control-display relationships on this scale; this ambiguity increases the danger of the control being turned in a direction which increases rather than decreases the error. Observations of transfer effects suggest that this danger is greater when the target position varies on the same or different instruments. If, in view of these considerations, the effective working area of the circular scale is reduced, the advantage of this scale in wrapping '10 in. of scale around a 3 in. dial' appears less impressive.

It is therefore suggested that although the three arrangements listed in group (A) above lead to an equally good mean performance, the horizontal linear scale may well be preferred, especially when erroneous control movements are to be avoided. It will be observed in this connexion that performance on this scale is more consistent than on any other.

Similar considerations apply to the position of the vertical scale in group (B).

It must be emphasized that the above conclusions are limited to tracking tasks in which the control knob is vertically below the display. It would be desirable to investigate a situation in which the knob is in some other position, for example, to the right of the display. In the case of the linear scales the optimum control-display relationships are probably the same, but it may well be found that performance on the vertical scale is now superior to that on the horizontal.

This article is based on reports submitted to the Flying Personnel Research Committee (Loveless, 1954*a*, *b*, 1955), for whom the work was carried out. The author is grateful to Prof. R. C. Browne for his encouragement, to D. J. Newell for his advice on statistical questions, and to the Officers Commanding Royal Air Force Stations, Acklington and Middleton St George, for their courtesy in making subjects available.

References

Carter, L. F. & Murray, N. L. (1947). A study of the most effective relationships between selected control and indicator movements. Chapter 10 in Fitts, P.M. (ed.), *Psychological Research on Equipment Design*. U.S. Govt. Printing Office.

Connell, S. C. & Grether, W. F. (1948). Psychological factors in check-reading similar instruments. *U.S.A.F. Air Materiel Command Mem. Rep.* MCREXD–694–7A.

Dallenbach, K. M. (1923). Position versus intensity as a determinant of clearness. *Amer. J. Psychol.* **34**, 282–6.

Fisher, R. A. & Yates, F. (1948). *Statistical tables for Biological, Agricultural and Medical Research*. Edinburgh: Oliver and Boyd.

Fitts, P. M. & Simon, C. W. (1952). The arrangements of instruments, the distance between instruments, and the position of instrument pointers as determinants of performance in an eye-hand co-ordination task. *U.S.A.F. Air Materiel Command Tech. Rep.* no. 5832.

Fitzwater, J. T. (1948). A study of the effects of rest pauses in perceptual-motor learning involving compensatory pursuit. Master's thesis, Ohio State University. Unpublished.

Gardner, J. F. (1950). Direction of pointer motion in relation to movement of flight controls. Cross-pointer type instrument. *U.S.A.F. Air Materiel Command Tech. Rep.* no. 6061.

Graham, N. E. (1952). Manual tracking on a horizontal scale and in four quadrants of a circular scale. *Brit. J. Psychol.* **43**, 70–7.

Grether, W. C. (1947). Direction of control in relation to indicator movement in one-dimensional tracking. *U.S.A.F. Air Materiel Command, Mem. Rep.* TSEAA–8–WGF.

Loveless, N. E. (1954*a*). Manual tracking in the four quadrants of a circular scale. II. Upper and lower quadrants. *F.P.R.C. Rep.* no. 867.

Loveless, N. E. (1954*b*). Manual tracking in the four quadrants of a circular scale. III. Left and right quadrants. *F.P.R.C. Rep.* no. 886.

Loveless, N. E. (1955). Manual tracking on horizontal and vertical linear scales. *F.P.R.C. Rep.* no. 914.

Mitchell, M. J. H. & Vince, M. A. (1951). The direction of movement of machine controls. *Quart. J. Exp. Psychol.* **3**, 24–35.

Murrell, K. F. H. (1951). The design of dials and indicators. *Naval Motion Study Unit Rep.* no. 48.

Murrell, K. F. H. (1952). Direction of motion relationships between valves and gauges when mounted in panels. Results from a printed test. *Naval Motion Study Unit Rep.* no. 51.

Simon, C. W. (1951). Effect of stress on performance with preferred and non-preferred instrument designs. *Amer. Psychol.* **6**, 387.

Simon, C. W. (1952). Instrument-control configurations affecting performance in a compensatory pursuit task. *U.S.A.F., Wright Air Development Center, Dayton, Ohio, Tech. Rep.* no. 6015.

Warrick, M. J. (1947). Direction of movement in the use of control knobs to position visual indicators. Chapter 9 in Fitts, P. M. (ed.) *Psychological Research on Equipment Design.* U.S. Govt. Printing Office.

Warrick, M. J. (1948). Direction of motion stereotypes in positioning a visual indicator by use of a control knob. II. Results from a printed test. *U.S.A.F. Air Materiel Command, Mem. Rep.* MCREXD–694–19A.

Warrick, M. J. (1949). Effects of motion relationships on speed of positioning visual indicators by rotary control knobs. *U.S.A.F. Air Materiel Command, Tech. Rep.* no. 5182.

White, W. J. (1949). The effect of dial diameter on ocular movements and speed and accuracy of check-reading groups of simulated engine instruments. *U.S.A.F. Air Materiel Command, Tech. Rep.* no. 5826.

(*Manuscript received* 16 *June* 1955)

SOME STUDIES OF PERCEPTION IN RELATION TO AGE

By JEAN G. WALLACE

Nuffield Research Unit into Problems of Ageing, Psychological Laboratory, Cambridge

A series of experiments to discover the extent of age differences in visual perception when temporal integration was required suggested that:

(1) Where the displays were simple, the performance of different age groups was nearly equal, but with increasing complexity of displays the older subjects were relatively increasingly slow in making correct identifications.

(2) With the complex displays there was a relationship between the degree of correct identification and (*a*) the cumulative viewing time; (*b*) the amount of the display seen at one moment, such that an increase in both led to a greater improvement in the performance of the older subjects than an increase in either time or amount alone. Given a sufficient increase, the performance of older subjects equalled that of younger.

(3) The greater difficulty of older subjects with complex material appeared to be due mainly to the greater amount of temporal integration required with such material.

I. Introduction

Previous studies of perceptuo-motor performance in relation to age have suggested that perceptual rather than motor factors often set the limits to performance among older people, and that this limit is not due to falling efficiency of the peripheral sense organs so much as to central changes (Welford *et al.* 1951). An experiment by Bernardelli (see Welford *et al.* 1951) suggested that one of the difficulties for older people lies in the relating together of information arriving serially in time—a finding which would seem to be in line with much clinical evidence about senile conditions, although not commonly recognized as a feature of normal middle or early old age.

Accordingly, displays were presented in such a way that only a small portion was visible at any one instant, but the whole was successively revealed. Correct identification of the displays depended upon integrating material perceived serially over a period of 2–3 sec.

The apparatus consisted of a box containing a funnel 15 in. long and 6 in. deep into which the subject looked through an aperture 5 in. across and $1\frac{1}{2}$ in. deep. At the end of the funnel was a horizontal slit, which could be varied in depth, directly in front of, and the width of, a well-illuminated band. The displays were attached to the band and seen through a transparent screen, and the band moved downwards behind the slit, the lowest portion of the displays being thus the first to appear.

The displays to be identified were boldly drawn in black ink on a white ground. They had previously been used in experiments by Dr M. D. Vernon. The subject was given warning of the approach of the display by two lines on the band which formed a Λ and appeared in the slit as two black spots which converged gradually towards the centre, reaching it just before the display appeared.

It was not possible to test the subjects' vision before the experiments, but all subjects who normally used glasses wore them, and subjects were questioned to ensure that they could see the displays properly.

Subjects were told to report verbally what they had seen, in as much detail as possible, immediately after each presentation.

II. Procedure and results

A. *Varying complexity in relation to age*

Experiment I

In this experiment a comparison was made of the performance of two age groups with displays having various degrees of complexity. The groups consisted of sixteen subjects aged under 30 and sixteen subjects aged over 60. The younger group were all naval ratings with an average age of 19·7 years. The older were mostly members of an Over Sixties Club, of both sexes, with an average age of 71·7.

Fig. 1. Examples of the four types of display used in Exp. I.

The displays (examples of which are shown in Fig. 1) consisted of four of each four types:

(*a*) outlines of geometric figures;

(*b*) silhouettes of simple figures (Silhouette I);

(*c*) 'meaningful' silhouettes (Silhouette II);

(*d*) more complex pictorial representations.

Two of each of the above types were shown under two different conditions, at a band speed of 1 in. per sec.:

(1) Each picture was shown first with the horizontal slit 0·1 in. deep. If the subject

failed to identify the picture correctly after one presentation at 0·1 in. he was shown it once at 0·2 in. and so on to 0·5 in., or to the point of identification if sooner.

(2) Each picture was shown at one aperture throughout—half of them at 0·1 in. and half at 0·3 in.—up to 3 times or until correct identification, whichever came sooner. Pictures were shown alternately at 0·1 and 0·3 in. for each subject. This condition was introduced as a control in order to discover how much improvement in recognition could be brought about merely by repetition of the display without widening of the aperture. The results will be given later when a number of controls are being discussed.

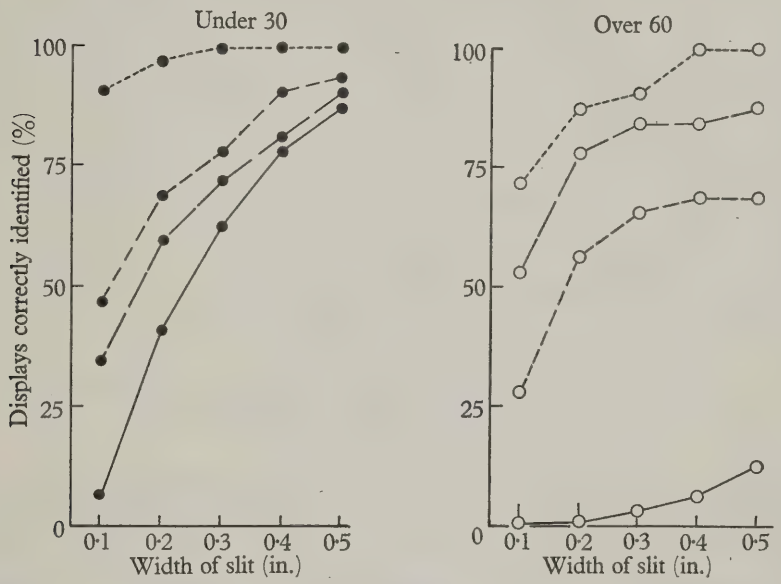


Fig. 2. Progress of correct identification as aperture increased. Exp. I. *Key*: Geometric Figures, dotted line; Silhouette I, large dashes; Silhouette II, small dashes; Pictorial Representations, black line.

Each subject saw eight of the pictures under one condition and eight under the other, and the pictures were seen an equal number of times under each condition by each age group. The order of presentation was varied within each group.

Results with increasing aperture (*Condition* 1)

The numbers of correct identifications following each presentation are shown in Fig. 2. The two age groups gave very similar results for the geometric figures. The older were a little less successful than the younger in identifying the silhouettes, and strikingly less successful with the pictorial representations. The differences in number of successes with age are shown in Table 1. There thus appears a transition from near equality between the age groups with the easiest material to wide divergence with the more difficult. This is in line with the results of several perceptuo-motor experiments where a similar transition has been shown (Szafran in Welford, 1951; Kay, 1954).

The possibility was investigated that the older subjects might be having greater difficulty because of inability to identify the displays correctly even outside the apparatus. No such inability was found, and even in the apparatus at the widest aperture possible ($1\frac{4}{10}$ in.), which was not wide enough to show any picture in entirety at one moment, failures to identify completely on one presentation only amounted to 17%.

Table 1. *Number of successes and analysis of failures, after five presentations. Exp. I, Increasing Aperture*

Type of display	No. of successes (max. 32)	No. of failures 1	2	No. failing
	Under 30			
Geom. Figs.	32	.	.	0
Sil. I	30	2	.	2
Sil. II	29	3	.	3
Pict. Repns.	28	4	.	4
				9
	Over 60			
Geom. Figs.	32	.	.	0
Sil I	22	6	2	8
Sil. II	28	4	.	4
Pict. Repns.	5	3	12	15
				41

Comparison of successes of under thirties and over sixties. (τ)*

Geom. Figs.	No correlation
Sil. I	N.D. = 2·09, $P = <0{\cdot}04$
Sil. II	No correlation
Pict. Repns.	N.D. = 4·35, $P = <0{\cdot}001$

* See Kendall, (1948).

B. *Course of age change in perception of complex material and its relation to occupational level*

Experiment II

In this experiment the aim was to investigate in greater detail the course of the age change in the perception of complex material and the relation of this to occupational level. To shorten the experiment only 'pictorial representation' displays were shown—the four used in Exp. I and four others of the same type.

Each subject saw half the pictures under one condition, half under another, the two conditions being:

(1) As in Exp. I, starting at 0·1 in. for the first presentation and rising by 0·1 in. on each subsequent presentation up to 0·5 in. or to correct identification if sooner.

(2) At a fixed aperture of 0·5 in. up to five presentations or correct identification if sooner. The second condition was again a control, the results from which will be considered later.

The subjects were divided into two groups according to occupational level. Group 1 consisted of naval ratings, men employed at a Government Training Centre and pensioners of a factory employing mainly unskilled and semi-skilled workers, and were comparable in level to the subjects of Exp. I. Group 2 consisted of University lecturers and research workers, those in administrative positions in industry and others between these two levels.

Each group consisted of eight subjects in each decade from the twenties to the sixties. There were also eight in group 1 in their seventies whose results are included wherever they do not affect comparison as between equal numbers of subjects.

Results with increasing aperture (Condition 1)

Numbers of correct identifications are set out for each age decade and for each occupational group in Fig. 3. The results are not entirely regular, but it can be seen that broadly speaking there is little difference between the twenties, thirties and forties in either

occupational group, but a substantial and significant difference between these and the subjects over 50. The differences in successes are shown in Table 2.

Table 2. *Number of successes and analysis of failures after five presentations. Exp. II, Increasing Aperture*

Age	No. successes (max. 32)	No. failures 1	2	3	4	No. failing (max. 8)
		Group 1				
21–30	30	.	1	.	.	1
31–40	20	3	.	3	.	6
41–50	23	1	1	2	.	4
51–60	17	3	1	2	1	7
61–70	10	1	1	5	1	8
71–80	9	1	3	.	4	8
		Group 2				
21–30	26	4	1	.	.	5
31–40	26	1	1	1	.	3
41–50	31	1	.	.	.	1
51–60	26	6	.	.	.	6
61–70	10	1	1	5	1	8

Comparison between number of successes of over and under fifties (τ)*

Group 1 (without 70's):
N.D. = 4·83, P = 0·001.

Group 2:
N.D. = 5·15, P = 0·001

* See Kendall (1948).

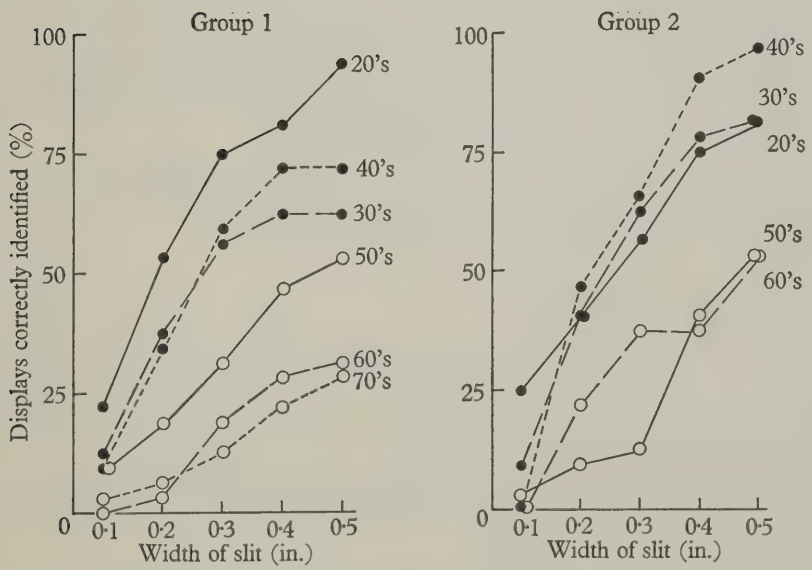


Fig. 3. Progress of correct identification as aperture increased. Exp. II.

The oldest subjects do rather better than those in Exp. I, and the younger subjects do so also on the first presentation. This is probably due to the displays presented in Exp. II being all of one type (i.e. pictorial representations), which meant that the subjects were able to approach the presentation of each new display with a smaller range of possibilities in mind.

The differences between the two occupational groups were small, but, except in the twenties, numbers of successes were in each age range greater among the subjects of higher occupational level.

The content of all the responses was analysed to discover in particular whether failures at identification by older subjects was due to an immediate, but incorrect, interpretation of the picture or to not giving any interpretation. Results are given in Table 3. It was

Table 3. *Type of incorrect or incomplete response given. Expt. II, Increasing Aperture*

Age	No idea	%	Indefinite	%	Definite	%	Total
			Group 1				
21–30	0	0	2	3·6	53	96·4	55
31–40	1	1·2	2	2·4	83	96·4	86
41–50	2	2·5	6	7·4	73	90·1	81
51–60	4	3·6	11	9·9	96	86·5	111
61–70	26	19·4	11	8·2	97	72·4	134
71–80	27	19·7	22	16·1	88	64·2	137
			Group 2				
21–30	2	2·8	8	11·3	61	85·9	71
31–40	5	6·8	2	2·7	66	90·5	73
41–50	4	6·2	3	4·7	57	89·1	64
51–60	7	5·8	15	12·5	98	81·7	120
61–70	11	9·8	12	10·7	89	79·5	112

found that there was a slight rise over 50 and a greater one over 60 in the percentage of responses which were either vague and unstructured or where no attempt was made at any interpretation. But the percentage of responses which involved some definite, concrete interpretation of the display never dropped below 64% (the seventies). Beyond such an analysis as this, it was not possible to gain any further information from the responses themselves as to the nature of the difficulty of the older subjects, except as regards repetition of errors, which will be mentioned later.

Using a less strict criterion for correct identification it was found that the age differences remained the same relative to each other, but that the stricter criterion was more useful for a more detailed investigation of the differences.

The nature of the difficulty involved

The method used in Exps. I and II of presenting a display several times at gradually increasing aperture has the disadvantage that several variables in the display conditions are changed at the same time, and it is not immediately possible to say which is the cause of any difficulty shown by the subjects.

In particular, where the number of displays correctly identified rises with successive presentations, the increase might at first sight be due to any one of three factors:

(*a*) Mere repetition of the exposure.

(*b*) The fact that as the slit widens each part of the display is exposed for a longer time.

(*c*) The fact that widening of the slit results in a larger section of the display being observable at any one time.

The psychological significance of (*b*) follows from the postulate by Hick (1952) that the rate of gain of information is constant, thus relating time to difficulty of identification. The significance of (*c*) is twofold. In the first place, the better articulation of the display as the slit is widened (and more is seen at once) will reduce the amount of information

that has to be extracted from the display for identification to be achieved. Secondly, when the aperture is made very wide, or the whole picture is seen at once, there is no need for temporal integration of the material viewed as it passes the slit.

C. *Effects of repetition*

Results with fixed apertures in Experiments I and II

The numbers of correct identifications for the exposures at 0·1 and 0·3 in. in Exp. I are set out in Table 4, and for the exposures at 0·5 in. in Exp. II in Table 5. Comparing these with the corresponding results for increasing aperture it can be seen at once that repetition

Table 4. *Number of correct identifications at each presentation. Expt. I*

	Fixed aperture 0·1 in.				Fixed aperture 0·3 in.			
	1	2	3	No. successes (max. 16)	1	2	3	No. successes (max. 16)
				Under 30				
Geom. Figs.	15	0	0	15	16	0	0	16
Sil. I	10	2	0	12	14	0	0	14
Sil. II	7	2	0	9	13	1	1	15
Pict. Repns.	1	0	2	3	6	4	0	10
Total				39				55
				Over 60				
Geom. Figs.	12	2	0	14	15	0	0	15
Sil. I	7	2	0	9	11	1	1	13
Sil. II	10	4	0	14	11	2	0	13
Pict. Repns.	0	0	0	0	2	0	0	2
Total				37				43

Table 5. *Number of correct identifications at each presentation. Exp. II, fixed Aperture 0·5 in.*

	Group 1						Group 2					
	Presentation					No. successes	Presentation					No. successes
Age	1	2	3	4	5	(max. 32)	1	2	3	4	5	(max. 32)
21–30	25	3	1	0	2	31	26	5	1	0	0	32
31–40	21	5	2	0	1	29	19	7	4	1	0	31
41–50	22	3	4	1	0	30	24	5	1	0	0	30
51–60	12	10	2	1	1	26	17	4	2	2	1	26
61–70	6	6	3	3	1	19	11	3	3	0	1	18
71–80	4	6	2	1	2	15	.	.	.	.	.	.

alone is not a sufficient explanation of the rise in the number of correct identifications under increasing aperture conditions. If it were, we should expect the first exposures under fixed aperture and increasing aperture conditions to be the same, and that the two types of condition would keep in step with the second and subsequent exposures. Clearly this does not happen: the numbers of correct recognitions at 0·1 in. after the first are too low, while the numbers of correct recognitions on the early exposures at 0·3 and 0·5 in. are too high.

D. *Effects of exposure time and aperture*

Experiment III. Exposure time reduced to one third that of Experiments I and II

It seemed possible that correct recognition might, to some extent, be a function of the total length of time for which a picture had been seen in each exposure together with any

previous exposures. Thus two exposures at 0·5 in. sec. (fixed speed) might be equivalent to a single exposure at 1·0 sec., or the first two exposures at increasing speed—0·1 and 0·2 in.—might together equal a single exposure at 0·3 in. Fig. 4 shows the results for the thirties of group 2 in Exp. II plotted with these cumulative exposure times on the abscissa. On the same graph are plotted also the results from another group of eight subjects of the same age range and occupational grade who repeated Exp. II but with the band moving at three times its previous speed (i.e. 3 in. per sec.)

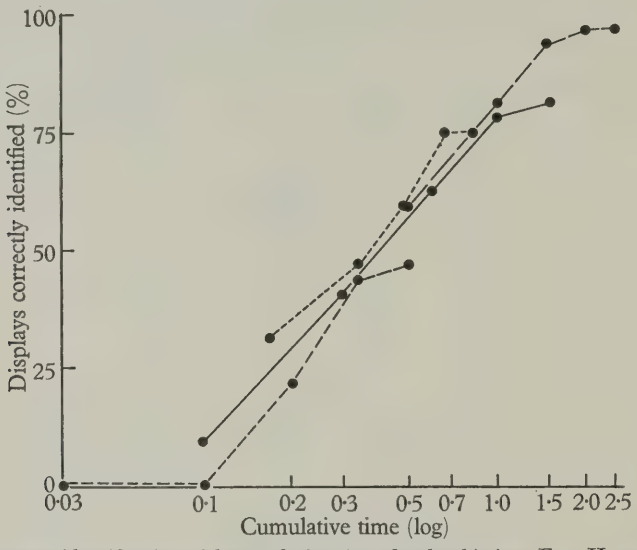


Fig. 4. Progress of correct identification with cumulative time, for the thirties. Exp. II and III. *Key*: Exp. II. Increasing aperture, black line; Fixed aperture 0·5 in., large dashes; Exp. III. Increasing aperture, small dashes; Fixed aperture 0·5 in., dotted line.

If cumulative exposure time were the only important factor the results should all lie along a single line, and there is in fact a close linear relationship between the number correctly identified and the cumulative (log) time, though the results with the 0·5 in. fixed aperture are consistently a little better than those with increasing aperture. It is interesting to note that the relatively high speed of *movement* of the band used in this experiment did not appear to have an adverse effect on performance.

Experiment IV. Increasing exposure time with aperture constant

Fig. 5 shows the results for the youngest and oldest age ranges of group 1 in Exp. II plotted in a manner similar to Fig. 4. Two other sets of results are plotted on the same graph. The first (Exp. IV) is from a group of eight naval ratings aged under 30 and a group of eight members of an Over Sixties Club with an average age of 69·5 who were given the same eight pictures as those used in Exp. II under the following conditions:

(*a*) Four with an aperture of 0·3 in. and the same speed as that used in Exp. I and II (1 in. per sec.), so that any part of the picture was visible for $\frac{3}{10}$ sec.

(*b*) Four with an aperture of 0·3 in. and the speed changing so that the first presentation was such that any part was visible for $\frac{1}{10}$ sec. (equal to the time for which it would be seen with an aperture of 0·1 in. at the speed of 1 in. per sec.), the second presentation so

that any part was visible for $\frac{2}{10}$ sec. and so with decreasing speed until it was visible for $\frac{5}{10}$ sec., or until correct identification if sooner. Condition (*b*) thus reproduced the exposure *times* for the increasing aperture condition of Exp. II while keeping the *aperture* constant.

Among the younger subjects the results with 0·5 in. aperture in Exp. II and with both conditions in Exp. IV fit the cumulative time hypothesis reasonably well, but all are superior to the increasing aperture condition of Exp. II. It seems as if the wider apertures of 0·3 and 0·5 in. confer an advantage during the first one or two exposures which persists to subsequent exposures.

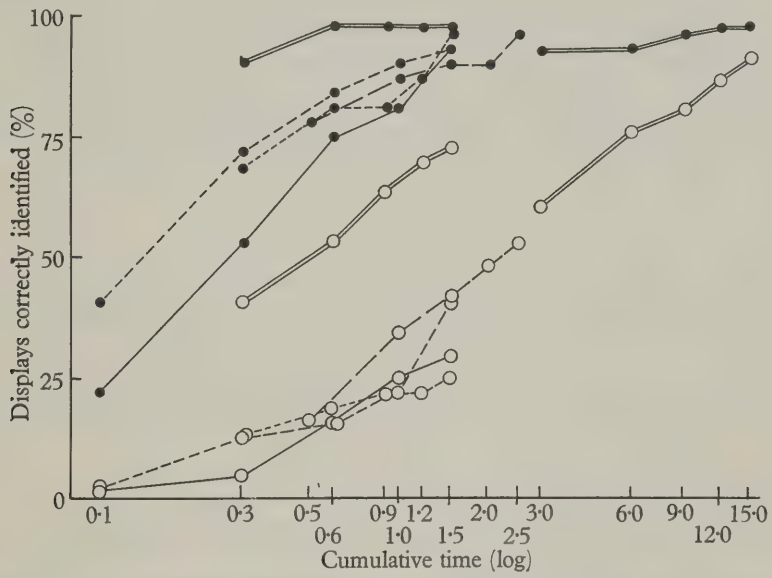


Fig. 5. Progress of correct identification with cumulative time, comparing oldest and youngest groups where aperture and speed conditions varied. Exps. II, IV and V. *Key*: Youngest group, closed circles; Oldest group, open circles. Exp. II. Increasing aperture, black line; fixed aperture 0·5 in., large dashes; Exp. IV. Fixed speed (0·3 in.), dotted line; increasing speed (0·3 in.), small dashes. Exp. V. Tachistoscope, open line.

The older subjects show similar trends although their results are less regular. The results with 0·5 in. aperture from Exp. II are superior to all the others, but the results from Exp. IV are not clearly different from those of the increasing aperture conditions of Exp. II.

It would seem that aperture has some effect over and above the exposure time with which it is confounded in the increasing aperture conditions.

Experiment V. Whole picture shown at once

Confirmation of this last point is obtained from the results of an experiment in which the same eight pictures were exposed in a tachistoscope so that the whole picture was visible at one time for $\frac{3}{10}$ sec. (i.e. the same length of time as was taken for each part of the picture to pass the aperture in condition (*a*) of Exp. IV). The subjects consisted of eight naval ratings with an average age of 19 and eight men and women of a comparable socio-economic level who were all in their seventies except one, who was 67 (average age 73).

The procedure and criteria for correct identification was as before, all eight pictures being exposed five times unless correctly identified sooner.

The results are shown in Fig. 5. The superiority of them to those obtained in the previous experiments is clear, especially among the older subjects.

Two other groups of similar ages and grades carried out the same experiment except that the exposure time was 3 sec., i.e. the time taken for the whole picture to pass the aperture at the basic band speed in the previous experiments (1 in. per sec.).

The results for these groups are shown at the right-hand side of Fig. 5. Correct identifications by the younger subjects are nearly complete on the first exposure. The older subjects during the first two exposures show results comparable with the tachistoscopic presentation at $\frac{3}{10}$ sec., but their correct identifications gradually climb to near the maximum possible by the fifth exposure, where the level they reach (i.e. after a cumulative exposure time of 15 sec.) is about the same as the younger reach at their first exposure at $\frac{3}{10}$ sec. Thus, although if they are given sufficient exposure time the older subjects reach the level of the younger, the extra time required is remarkably long.

Experiment VI. Varying aperture with exposure time constant

Further evidence of the importance of aperture as opposed to exposure time is furnished by an experiment in which aperture and band-speed were varied together so that the exposure time was kept constant at $\frac{3}{10}$ sec. whatever the aperture.

A group of eight naval ratings under 30 and a group of eight subjects from an Over Sixties Club (average age 69·9) were shown the same eight pictures under the following conditions:

(*a*) Four of the pictures with the first presentation at 0·1 in., the next at 0·2 in. and so on to 0·5 in. or correct identification.

(*b*) Four of the pictures with the first presentation at 0·5 in., the second at 0·4 in. and so on to 0·1 in. or correct identification.

The results are shown in Fig. 6(*a*), and it can be seen clearly that the subjects do better when the wide apertures are presented first. The fact that the superiority of the wide apertures persists until the fifth exposure indicates that, as in Exp. IV, the wider apertures confer an initial advantage which persists.

The experiment was repeated with a further sixteen subjects of an occupational level similar to group 2 in Exp. II: eight with ages ranging from 25 to 38 (mean 29·8) and eight from 42 to 54 (mean 46·6). The results from this second group are shown in Fig. 6(*b*) and were consistent with those of the former groups.

Repetition of responses

Some indication of one possible reason for incorrect identifications by older people is given by comparing successive replies. In the course of Exp. I there was evidence of a 'set' being formed by older subjects, the contents (or supposed contents) of one display influencing the response to the next. Seven of the older subjects showed this tendency, several of them more than once, and only three of the younger subjects did this, in no case as clearly. For example, an older subject gave this series of responses; to the arrow (Sil. I) 'shaped like a fan', followed immediately as a response to the star (Sil. I) by 'same as the fan but opens out like a tent' and then for the next stimulus, the triangle (Geom.

Figs.), 'a tent—a triangle'. The same subject, with the series at fixed aperture, gave these responses: having identified the yacht (Sil. II) correctly, responded to the monk (Sil. II) as 'someone signalling on a steamer', and to the next stimulus, a man carrying a pile of records (Pict. Rep.), 'something on the sea, a lighthouse'.

In a similar way, in this experiment, older subjects tended more frequently than younger to repeat errors made on one presentation of a picture when responding to a subsequent presentation of the same picture. The number of exact repetitions and the

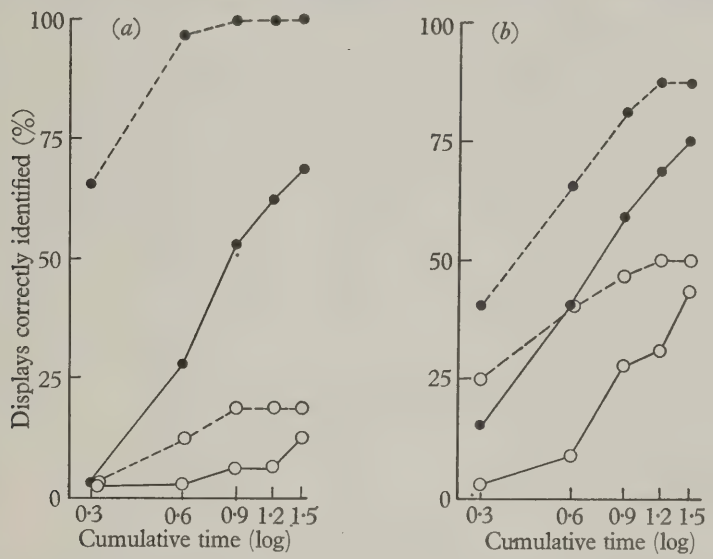


Fig. 6. Progress of correct identification with cumulative time, where exposure time was constant and aperture varied. Exp. VI: (*a*) group 1, oldest and youngest subjects; (*b*) group 2, intermediate age groups. *Key*: Younger subjects, closed circles; older subjects, open circles; decreasing aperture, dotted line; increasing aperture, black line.

percentage they formed of the total errors are given in Table 6(*a*) for all the experiments. It will be seen that a rise in repetition with age occurs often although not in all cases. In Exp. II, consideration was given to responses which, though not exact repetitions of previous ones, were very nearly so—the general content being the same and any changes being very small. Table 6(*b*) shows that there was a consistent increase with age in such repetitions where the aperture was fixed at 0·5 in.

Throughout the experiments, responses which contained material partly or wholly extraneous to the actual display were repeated more often than those correct but insufficient, but there was no consistent age difference here.

The tendencies to repetition are similar to those noted by Verville & Cameron (1946) in an experiment in which adults of different ages had to identify incomplete pictures, and in a learning experiment by Kay (1951). Verville & Cameron found that one picture was frequently wrongly interpreted as an animal, and older subjects perceiving the figure in this way persistently named animal after animal, instead of shifting to a different object when they found they were wrong. The fact that these tendencies appear among younger as well as older subjects in our experiments makes it seem probable not so much that they are a special feature of old age as that changes occur which may, in certain circumstances, throw them into increasing prominence as age advances.

Table 6. (*a*) *Exact repetition of responses, other than 'No Idea'*

	No. of repetitions	No. subjects repeating at least once	Av. repetitions per subject	Repetitions as percentage opportunity to repeat
		Increasing aperture or speed		
Exp. I				
Under 30	17	10	1·7	13·0
Over 60	43	14	3·1	20·9
Exp. II				
Group 1				
21–30	3	3	1·0	5·7
31–40	15	6	2·5	20·3
41–50	5	4	1·3	6·9
51–60	11	5	2·2	11·5
61–70	13	6	2·2	13·3
71–80	12	5	2·4	10·5
Group 2				
21–30	3	2	1·5	4·6
31–40	5	4	1·3	7·5
41–50	1	1	1·0	1·6
51–60	4	4	1·0	3·9
61–70	8	6	1·3	8·6
Exp. IV				
Under 30	5	3	1·7	13·9
Over 60	18	6	3·0	16·1
		Increasing aperture and speed		
Exp. VI				
Under 30	13	7	1·9	16·1
Group 2, av. 29·8	5	3	1·7	7·2
Group 2, av. 46·6	14	5	2·8	13·3
Over 60	15	8	1·9	12·3
		Fixed aperture or speed		
Exp. I				
Under 30	15	11	1·4	15·6
Over 60	31	13	2·4	32·6
Exp. II				
Group 1:				
21–30	1	1	1·0	5·9
31–40	5	3	1·7	20·0
41–50	3	2	1·5	13·6
51–60	9	3	3·0	20·0
61–70	29	6	4·8	31·9
71–80	24	7	3·4	26·9
Group 2				
21–30	0	0	0	0
31–40	2	1	2·0	9·1
41–50	1	1	1·0	6·7
51–60	7	3	2·3	14·9
61–70	20	5	4·0	27·4
Exp. IV				
Under 30	8	3	2·7	30·8
Over 60	36	8	4·5	34·6
		Decreasing aperture and speed		
Exp. VI				
Under 30	0	0	0	0
Group 2, av. 29·8	14	4	3·5	35·0
Group 2, av. 46·6	28	6	4·7	36·8
Over 60	26	8	3·2	23·4
		Fixed aperture and speed		
Exp. V (tachistoscope)				
Under 30	2	1	2·0	22·2
Over 60	21	8	2·6	10·9

Note. 'Number of repetitions' alone is misleading because the better the performance the less the opportunity for repetition, but greater opportunity to repeat (because of incorrect identification) may be due partly to repetition. Thus, the first and last columns in this Table should be considered together.

Table 6(*b*). *Experiment II. Exact plus very nearly exact repetitions of responses, other than 'No Idea'*

	No. of repetitions	No. subjects repeating at least once	Av. repetitions per subject	Repetitions as percentage opportunity to repeat
		Increasing aperture		
		Group 1		
21–30	10	5	2·0	18·9
31–40	28	6	4·7	37·8
41–50	16	5	3·2	22·2
51–60	30	7	4·3	31·2
61–70	20	8	2·5	20·4
		Group 2		
21–30	7	5	1·4	10·8
31–40	13	4	3·2	19·4
41–50	6	4	1·5	9·5
51–60	21	8	2·6	20·8
61–70	20	6	3·3	21·5
		Fixed aperture, 0·5 in.		
		Group 1		
21–30	3	1	3·0	17·6
31–40	7	4	1·7	28·0
41–50	8	4	2·0	36·3
51–60	19	6	3·2	42·2
61–70	43	6	7·2	47·2
		Group 2		
21–30	1	1	1·0	16·7
31–40	5	3	1·7	22·7
41–50	4	1	4·0	26·7
51–60	15	3	5·0	31·9
61–70	29	6	4·8	39·7

III. Discussion

From the practical standpoint the indications of these experiments are fairly straightforward. Identification by older subjects of complex displays seen with only a part visible at once or for a restricted time, was much less accurate than it was by younger subjects. When the displays were simple the achievement of older subjects almost equalled that of younger. When they could be seen whole and for a longer time differences in performance with age became much smaller. Simplicity of display, full visibility and ample viewing time would, therefore, seem important conditions of success for older people.

Two sorts of differences have emerged therefore between older and younger subjects:

(1) *Where complexity of display varied*

It is suggested that, under conditions such as in Exp. I, the greater the complexity the greater the amount of temporarily meaningless material which has to be carried and integrated to produce correct identification. If it were simply the greater amount of material to be apprehended at once which was the source of difficulty older subjects should perform worse the wider the aperture, and worst of all under tachistoscopic conditions, and this does not happen. If temporal integration of *any* amount of material was equally difficult for older subjects their performance should not be worse with increasing complexity. It should be noted that meaningfulness of material apparently offset complexity for older subjects when complexity was not very great (Sil. II).

(2) *Where complexity was equal, but amount seen and length of viewing time varied*

Older subjects required either to see more of the display at once or a longer viewing time or both for their performances to equal those of younger subjects. Several possibilities can be suggested for this fact:

(*a*) *Directly physiological factors*, such as changes in the eye or in the central mechanism. The experimental results enable us to exclude one possibility under this head, namely, visual acuity as measured by test types with no account taken of reading time. Acuity in this sense of the term could not have been the cause of the observed age difference on performance because both young and old subjects were able to identify the pictures out of the apparatus and in the tachistoscope given sufficient time. Changes with age in the *speed* of dealing with fine detail in a display, as demonstrated by Weston (1949) may, however, be important, although how far these changes are dependent on peripheral factors and how far on central is difficult to say. The relation between acuity as normally measured with untimed tests and speed of perception and visual performance generally would seem to merit further study especially in relation to age.

(*b*) *Attitudinal factors*, such as extra caution on the part of older subjects. There is clear evidence of extra care being taken by older subjects in some sensori-motor tasks (see, for example, Brown in Welford *et al.* 1951), but it is not possible to say whether this is basically an attitudinal matter borne of experience, or whether it represents a shift in the manner of performance in the face of physiological changes in the organism which set a premium on economy of action and thus produce greater accuracy.

The increase with age in the present experiments of the number of indefinite and 'no idea' replies would be consistent with a theory that the older subjects were being more cautious, but would also be consistent with other theories, and cannot therefore be taken to support an explanation in attitudinal terms.

(*c*) *Experience.* Three possibilities need to be considered under this head:

(i) Experience increasing with age may widen the range of possible identifications by older subjects and thus increase the task of arriving at any particular one. Hick's (1952) law relating choice reaction time to degree of choice would lead us to expect that with increasing range of experience the time taken to make any particular identification would become longer.

(ii) Even if the total range of experience remains the same, some parts of it may become weighted at the expense of other parts as a man gets older. It seems reasonable to suggest that the process is akin to that demonstrated during these experiments in the influence of 'set' on perception. The experiments of Crossman (1953) and of Hymen (1953) would lead us to expect that objects to be identified which fall within the weighted part would be identified quickly, but those falling outside it would take a relatively much longer time.

(iii) The tendency for repetition of errors and the effects of 'set' shown in some of the present experiments and in those of Verville & Cameron (1946) suggests that effects of previous experience may have been operating on a very short time scale from one response to the next. It is possible that a subject's own previous response to a display may form a barrier to a change of interpretation. The greater repetition of responses involving extraneous material supports this suggestion. Kay (1955) has demonstrated such an influence with verbal material. There is some indication that such a tendency increases with age.

These explanations might also account for some of the increased slowness of older subjects with increased complexity.

In general the results bear out the expectations with which the experiments were started, in that they do show well-marked similarities between changes with age in perception and in sensori-motor performances, such as age differences with increasing complexity of material. Moreover, these age differences were, at least in part, due to slowness at reaching the same attainment rather than inability to attain the same level. Another similarity lies in the tendency of the older subjects to repeat errors even in the face of information that they were wrong (cf. Kay, 1951). These similarities do not, of course, prove that age changes in perception and sensori-motor performance are caused in the same way, but they do indicate that they share common functional principles.

The author wishes to thank Mr A. T. Welford for much helpful criticism and advice throughout the experiments, and Dr J. Szafran and Dr H. Kay for much useful discussion.

REFERENCES

CROSSMAN, E. R. F. W. (1953). Entropy and choice time. *Quart. J. Exp. Psychol.* **5**, 41–51.

HICK, W. E. (1952). On the rate of gain of information. *Quart. J. Exp. Psychol.* **4**, 11–26.

HYMEN, R. (1953). Stimulus information as a determinant of reaction time. *J. Exp. Psychol.* **45**, 188–96.

KAY, H. (1951). Learning of a serial task by different age groups. *Quart. J. Exp. Psychol.* **3**, 166–83.

KAY, H. (1954). The effects of position in a display upon problem solving. *Quart. J. Exp. Psychol.* **6**, 155–69.

KAY, H. (1955). Learning and retaining verbal material. *Brit. J. Psychol.* **46**, 81–100.

KENDALL, M. G. (1948). *Rank Correlation Methods.* London: Griffin.

VERVILLE, E. & CAMERON, N. (1946). Age and sex differences in the perception of incomplete pictures by adults. *J. Genet. Psychol.* **68**, 149–57.

WELFORD, A. T. *et al.* (1951). *Skill and Age: An Experimental Approach.* Oxford University Press for the Nuffield Foundation.

WESTON, H. C. (1949). On age and illumination in relation to visual performance. *Trans. Illum. Engng Soc.* **14**, 281.

(*Manuscript received* 16 *June* 1955)

AN EXPERIMENTAL INQUIRY INTO THE CONCEPT OF PERCEPTUAL DEFENCE

By BERNARD R. SINGER

Birkbeck College, University of London

Experiments on the concept of perceptual defence can be considered as the result of two research interests: (1) attempts to relate experimental work to clinical findings and (2) attempts to determine the precise variables involved in the association between personality and perception. The concept of perceptual defence is based on the fact that subjects appear to show higher thresholds to taboo words and that autonomic responses to such words may be made before they are correctly recognized. This concept has been criticized on the grounds that the results may be more simply interpreted in terms of factors of word frequency and of suppression of responses by the subject. Other work has attempted to show that individual performance requires prime consideration.

This study has attempted to take account of these factors. A word-association test consisting of the most frequent words in the Thorndike and Lorge count was given to all subjects. On the basis of each subject's results, two lists were prepared for each subject, each list consisting of emotional and neutral words. A tachistoscopic presentation to the subject of one list was followed two weeks later by a presentation of the second list. It was found that six of the twelve subjects showed higher and two subjects lower thresholds to the emotional words on both occasions. The correlation between the two tests in terms of individual performance was highly significant.

Analysis of this and other experiments shows that the importance of factors of word frequency when familiar material is used remains to be determined. Further, some means of studying such factors other than by the use of word counts is necessary. It is suggested that a quantifiable 'hypothesis' or 'set' theory possesses the greatest advantages for an understanding of this area of research.

I. Introduction

The influence of past experience on perception has long been known in the study of human behaviour. The method by which such a process may operate has received several interpretations and awaits further clarification (see Bartlett, 1932; Oldfield, 1954). That the factors which affect perceptual processes may function without the conscious participation of the observer was known also to the earlier experimental psychologists (see Humphrey, 1951). Külpe (1904) and Jaensch (1930) further suggested that individual typologies may be related to characteristic forms of perception and imagery.

In keeping with the recent tendency to unify what were previously relatively separate fields of psychological study, attempts have been made to interpret current experiments on human perception in terms of theories of personality functioning, particularly that of Freud. These interpretations do not normally use the whole body of such theories, but attempt to apply certain specific principles, such as that of repression, to the experimental findings. This use of personality and psychodynamic theories has been especially congenial to those psychologists who have stressed the importance of internal, autistic factors as determinants of perception (Blake & Ramsey, 1951); in contrast, on the one hand, to those workers who place greater emphasis on the organizational properties of the perceptual field (Luchins, 1951), and on the other, to those who incorporate learning processes as the variables requiring prime consideration (Postman, 1953).

Stemming partly from Brunswik's functionalist interpretation of behaviour (Brunswik, 1955), a number of psychologists have attempted to examine the place of internal, or of

that which they term 'behavioural', determinants in perception (Postman, Bruner & Walk, 1951). Certain of these psychologists, particularly Bruner (1951*a*, *b*), have attempted more comprehensive conceptualizations to account for the experimental findings and to relate such work to clinical observation.

In the course of experimentation with the latter theoretical orientation as a background, certain observations were determined which possessed a striking similarity to one of the suggested psychoanalytic mechanisms of defence. Such findings, however, have resulted in differences of opinion as to whether the phenomenon of perceptual defence actually exists. Further, if such behaviour can be reliably determined from controlled experimentation, can it not be interpreted in a more parsimonious fashion? This particular situation has resulted in the performance of numerous experiments with the purpose of clarifying the issues involved. Unfortunately, however, such a situation has made it extremely difficult for anyone who wishes to determine the present position to do so without a large degree of labour and confusion from the apparently contradictory findings.

It is the purpose of this paper to present the situation as concisely as possible, and, by means of an experiment which has attempted to take account of the several issues and experimental findings involved, to assist in some measure in bringing clarification into this important area of research.

II. EXPERIMENTAL DEVELOPMENT AND PRESENT POSITION

(i) *Initial establishment of the concept*

Perhaps the first occasion on which the term perceptual defence was used in this context was that on which Bruner & Postman (1947) suggested such a term as a partial explanation of the results of an experiment they had undertaken. To describe this experiment briefly: ninety-nine words, consisting of taboo words and neutral words, were presented to nineteen subjects as a word-association test. For each subject, the six presented words giving the longest response time, the six words of shortest response time and the six words of medium time were extracted. A fortnight later, each subject was shown these eighteen words by means of a tachistoscope, and the exposure times necessary for correct identification of the individual words were determined.

Seventeen of the nineteen subjects showed a significant curvilinear relationship between their individual reaction times to words and the exposure times necessary for correct recognition of these words. The results were then examined for the subjects as a whole. Recognition scores and reaction times were standardized around each subject's own means, and the relation between average recognition times and reaction times was given a graphical treatment. The resulting curve was of an inverted U form, suggesting to the authors that certain subjects showed higher exposure times for words which gave a longer reaction time and that other subjects recognized such words more easily and thus behaved in the opposite sense. To interpret these results, the authors presumed that individual subjects have differing selective forms of dealing with emotional words. Those subjects who required longer exposure times to words whose emotional significance was indicated by length of reaction time appeared to indicate the operation of a form of defence mechanism called perceptual defence. Those subjects who, on the other hand, saw such words more readily, indicate the operation of a form of perceptual sensitization to potentially disturbing words.

It was McGinnies's (1949) experiment which aroused the most discussion. In this experiment, words were not initially given in the form of a free word-association test, but a short list of eighteen words was prepared consisting of seven socially taboo words such as 'raped', 'bitch' and 'whore', with the addition of eleven neutral words. This list was then presented as in the second part of the previously considered experiment to individual subjects, and the tachistoscopic exposure times were determined for the neutral and emotional words. Records were also taken of the psychogalvanic responses of the subjects before and after recognition of each word. The results are essentially as follows: (*a*) taboo words require exposures of a longer duration for correct recognition than the words which are considered to be neutral; (*b*) psychogalvanic responses after words have been exposed and not yet correctly identified are greater when the words not reported as seen are taboo than when they are neutral. McGinnies interpreted these results to suggest that there was a mechanism by which visual stimuli were 'filtered' in order to protect the individual as much as possible from being aware of words which have an unpleasant emotional connotation.

(ii) *Word frequency and response suppression*

This experiment was soon subjected to a detailed criticism by Howes & Solomon (1950). The main points of their criticism were, first, that the duration thresholds for words vary as a function of the logarithms of the relative frequencies of usage of these words, as determined by the Thorndike & Lorge (1944) word frequency counts, and that the exposure times for recognition of both types of words would thus be of the same order when the effects of word frequency are extracted (see McGinnies, 1950). Secondly, that the experimental atmosphere, together with the presence of Prof. McGinnies and his female assistant, would lead to the withholding of responses to taboo words until the subject was absolutely certain that such words were being exposed. The possible feelings of a subject in such a situation are most amusingly described by Howes & Solomon, but their argument must none the less be taken as a serious criticism of the conditions of McGinnies's experiment. The subject could conceivably then be seeing a taboo word, with a consequent deflexion of the psychogalvanometer, before the word had been correctly recognized in terms of the experimenter's criterion of a correct verbal report.

Further papers by these two authors (Howes & Solomon, 1951; Solomon & Howes, 1951) tended to substantiate their previous findings and to suggest that word frequency was an equally important variable in the study of individual values as measured by the Allport–Vernon test. The second suggestion, however, was criticized by Brown & Adams (1954) and by Dukes (1955).

The experiment of Bruner & Postman (1947) was repeated in several respects by Delucia & Stagner (1954) but with allowances made for word frequency. Their findings were in agreement with the early suggestion made by Bruner to McGinnies (see McGinnies, 1949), that both the factors of frequency and the emotionality of words were involved in such experiments.

To examine the part played by the subject's attitude and by response suppression in these experimental situations, Postman, Bronson & Gropper (1953) used four main subgroups which were each given a different set of instructions prior to presentation of the emotional and neutral words. One group, as in McGinnies's experiment, were not informed that taboo words would be shown. A second group was clearly instructed to

expect taboo words. A third group was not only warned to expect such words but was told that difficulties in recognizing taboo words was an indication of poor adjustment. A final subgroup was told that difficulties in perceiving taboo words was an indication of social adjustment. The resulting recognition thresholds were found to vary with the specific instructions given. The third group, for example, had the lowest thresholds for the taboo words and the first group showed the highest thresholds. It is significant that in this experiment, in which an attempt was made to control for word frequency, in the case of each subgroup, the thresholds for the taboo words were slightly lower than those for neutral words. The experimenters attempt to interpret this fact in terms of slight differences in frequency between the two sets of words as chosen for the experiment. Lacey, Lewinger & Adamson (1953), however, also noted that with an expectancy for taboo words, the threshold for such words was lower than for neutral words. Freeman (1954) and Bitterman & Kniffin (1953) also attempted to demonstrate that perceptual defence was not found when subjects were instructed to expect words of an emotional significance. The results of an experiment by Lawrence & Coles (1954) is interesting in that they showed that instructions were equally effective in reducing recognition thresholds when given either before or after the test flash. This finding would seem to indicate that it is the form of the subject's response which is affected by the conditions rather than his actual perceptions.

In two experiments carrying out substantially the same procedure (Solomon & Postman, 1952; King-Ellison & Jenkins, 1954), the frequency of syllables used was built into the actual experiment in order to study the relation between frequency and recognition thresholds. The subjects were initially presented with paralogs (pairs of nonsense syllables), certain of these paralogs being shown more often than others. A tachistoscopic presentation was then undertaken to determine the recognition thresholds for these paralogs. A high negative correlation was obtained between the frequency of initial presentation of the paralogs and their recognition thresholds, indicating that the less frequently a paralog was seen, the higher the threshold value necessary for correct identification.

(iii) *The clinical approach to the concept*

Two authors, Lazarus (1954) and Eriksen (1954), have undertaken a detailed criticism of the experimental work in this area. As these writers possess a similar point of view, it will be convenient to treat their opinions together. They agree with the objections of Solomon & Howes (1951) and of Postman *et al.* (1953) to the experiment of McGinnies, but feel that these authors have been rather selective in the experiments cited for criticism. In their opinion there are several experiments (to be treated shortly) which are not subject to such objection. Their own criticisms extend both to the McGinnies type of experiment and to that of Postman *et al.* (1953). These experiments, they consider, show two main defects. First, defence mechanisms from the point of view of Freudian theory may vary amongst individuals in their method of functioning. It is clear that there are other ways than repression of dealing with anxiety-arousing stimuli and that the treatment of data for groups as a whole would not show this. Secondly, if a form of perceptual defence is to be demonstrated, it is necessary to ensure that the words denoted as emotional are actually disturbing to the specific individuals. A particular word generally considered to be emotional may be in no wise disturbing to particular subjects.

Again, if frequency is a serious determinant of the phenomenon of perceptual defence, why is it that certain individuals recognize infrequent taboo words more readily? In the opinion of Lazarus, from an examination of the various experiments and of the data of Postman *et al.* (1953), word frequency is a serious factor only when highly infrequent words are used. From his point of view, word frequency may be considered equally well as a resultant of the personal needs and values of the whole population.

Lazarus and Eriksen cite experiments in which clinical typologies or the individual results of exposure to stresses have been found to correspond to the differential perception or recall of neutral and emotional words. Lazarus, Eriksen & Fonda (1951) demonstrated that good auditory perception of emotional material was characteristic of obsessive-compulsive patients. The reverse was found for hysterical personalities. Eriksen (1952) showed that subjects who had higher recognition thresholds and slower learning for words with a long delay than for words with a short delay in a word-association test, remembered more successes of a task. Those subjects who recalled more failures than successes learned words at a faster rate and had similar recognition thresholds for both short- and long-delay words in the association test. It is to be noted that in this experiment the subjects were instructed to expect possibly disturbing words. Finally, Lazarus & Longo (1953) showed that selective recall of failures in an interrupted task was associated with a relatively higher recall of nonsense syllables to which an electric shock had been paired. Those subjects who recalled a greater number of successes also recalled a larger number of syllables which had not been paired with the painful shock.

(iv) *Autonomic discrimination without awareness*

It has been mentioned above that in McGinnies's (1949) experiment, psychogalvanic responses were given to taboo words prior to verbal report. It was indicated that this may have been due to the possibility that the subjects were actually perceiving such words and yet not reporting them. Lazarus & McCleary (1951) undertook a most carefully controlled experiment to examine the possibility of discrimination without awareness. The complexity of design of this experiment does not allow a detailed report of their procedure and only the briefest description may be given. Ten nonsense syllables were presented separately to each of the subjects, five of these syllables being paired with an electric shock. When effective psychogalvanic responses had been conditioned to the shocked syllables, all ten syllables were presented at tachistoscopic exposure speeds too rapid for conscious discrimination (defined as the subject's inability to report which stimulus was presented). The syllables which had been paired with shock gave significantly greater psychogalvanic responses than the neutral syllables. Such autonomic discrimination without awareness is called 'subception' by the writers. As parts of this experiment have been misreported on occasion, it must be emphasized that there was no significant difference in the frequencies of identification of shocked and non-shocked syllables.

A number of writers consider that the subject obtains some information from the exposed syllable even though a correct verbal report is not given (Bricker & Chapanis, 1953). Howes (1954) has attempted a different account of discrimination without awareness by using a statistical theory which incorporates the subject's reports and which assumes the probability-like functioning of discrimination. Murdock (1954) has similarly subjected the interpretation of Lazarus & McCleary's experiment to some criticism.

(v) *Terminological aspects*

At this point the value of using a terminology derived from psychopathology must be touched upon. Klein (1954) objects to the immediate inclusion of psychoanalytic mechanisms as an explanation or description of the experimental results of such work. It is a presumption, he believes, to assume that if a form of defence is shown it must necessarily be the result of anxiety. The term 'defence' is useful when limited to forms of mastering anxiety, such as conflict among needs and the resulting avoidance of certain gratifications. Klein would prefer to use the term 'cognitive control' to indicate the particular observable form in which individuals manifest their internal processes in perception. It must be pointed out, however, that when the term 'perceptual defence' was first used by Bruner & Postman (1947), it was not intended to suggest that it necessarily corresponded to psychoanalytic processes but was used, as far as the present writer is aware, as a convenient descriptive term.

(vi) *Considerations for further experimentation*

The above broad treatment of the present position indicates that four factors in the design of such experiments require emphasis: (1) Factors of word frequency must be effectively controlled. (2) The data and the design of the experiment must be such that individual performance can be examined. (3) The words used must be of emotional significance to the individual subjects in the experiment. (4) The experimental conditions must be such that the subjects will not suppress their responses.

It has been the purpose of the writer of this paper to include these considerations in the design of the experiment to be presented in the following section.

III. EXPERIMENT

(i) *Selection of materials*

The essential difficulty in allowing for factors of word frequency was to obtain a frequency count which could be used satisfactorily with subjects who had spent their lives in Britain. Extensive inquiries failed to elicit the existence of such a frequency count of any comprehensiveness. The word counts available for use in this country, if not based upon American sources, were of such a nature that, although they have proved of value for language teaching, the wide coverage of source material essential for precise experimentation was lacking. Thus the only comprehensive source that could be used, if the experiment was to replicate to any extent the previous work in this field, was the Thorndike and Lorge *Word Book* (1944). It was decided to use the thousand most frequent words in the American language as indicated in this word count. The advantages of this particular selection appeared to be several. First, the sources from which these most frequent words were taken consisted of a relatively greater proportion of common British classical and children's books of the type with which the subjects are most likely to be acquainted. Secondly, the use of the most frequent words would meet the suggestion that frequency factors in relation to recognition thresholds are negligible when the words used are very common. The value of this point rests on the possibility that the most frequent American words in the count are not too dissimilar to the most frequent British words. Thirdly, the words used were of equivalent frequency and possible deviations from this range are likely to be minimal. Fourthly, it was considered more valuable to include words of

known equivalent frequency than to use either the methods which are often employed or to use those methods which, by the inclusion of certain taboo words of no reasonably known frequency, introduce an unnecessary degree of arbitrariness into the interpretation of the experimental findings. Finally, the use of words of such a common character greatly minimizes the possibility of the subjects suppressing their responses as a result of the experimental conditions.

In order thus to prepare a list of words for use in a word-association test and for later tachistoscopic presentation, the total number of five-lettered words (two hundred and four) were extracted from the thousand most frequent words in the Thorndike and Lorge word count. These words were then given separately to five male members of the department, who were asked to choose the fifty words from amongst the total number which they

Table 1. *List of forty-five 'emotional' and forty-five 'neutral' five-lettered words used in the word-association test**

about	since	north	*class*	*mount*
board	these	*under*	bring	reply
laugh	fruit	raise	built	going
enter	*alone*	sweet	*large*	*voice*
women	*small*	*young*	earth	*close*
blood	month	*storm*	*fight*	*first*
until	*court*	whole	those	*teach*
learn	*happy*	*truth*	white	clear
would	chair	guess	*sleep*	right
seven	prove	apple	*smoke*	*spend*
child	*marry*	south	among	paint
place	quite	still	along	*scene*
break	*death*	*dress*	*enemy*	sense
eight	*trust*	offer	heavy	*uncle*
cross	*money*	*dream*	*night*	ready
hurry	music	*shout*	*mouth*	*think*
tried	reach	*study*	*guard*	light
there	*power*	*drink*	third	today

* The words selected as emotional are in italics.

considered to be the most emotional in connotation, and the fifty words considered to be the most neutral. A final list was compiled consisting of forty-five emotional words and forty-five neutral words in the choice of which four of the five judges had agreed. This list was arranged in a random order and is shown in Table 1.

(ii) *Word-association test*

The subjects were twelve in number, all of whom were female, and consisted of four psychology undergraduates and eight other undergraduates or girls of the administrative staff. All the subjects were acquainted and on friendly terms with the experimenter. None of the subjects was informed as to the purpose of the experiment. On questioning the subjects when the full experiment was completed, it was determined that none of the subjects possessed any idea of its real purpose. When initially asked to participate in the experiment, they were told that the aim of the experiment was a general investigation into the 'seeing' of words. In the course of the questioning, two of the psychology undergraduates professed a minor acquaintance with the experiments of McGinnies (1949) when mentioned, but were vague as to its findings and did not at all relate it to the present experiment. These two subjects gave results contrary to each other when acting as subjects in the experiment.

The prepared list of ninety words was presented on separate occasions to each of the twelve subjects as a free word-association test. The following instructions were read to the subject at the commencement of the test:

I will call out a word and I would like you to reply with the first word which comes into your mind. It needn't be connected with the word I am giving; I just want you to say as quickly as you can the first word which pops into your head. I am not going to write down your replies. All that this clock does is to tell me the time it takes, which is all in which I am interested. The word given will only be said once. I will try to say it as clearly as I can, so please don't ask me to repeat it.

An electronic 'Labgear' timer (Type D 4102) using Dekatron glow-transfer tubes was started when the experimenter (male) called out a word and simultaneously pressed a key. A voice-key was used such that the subject's response stopped the activity of the electronic timer, indicating the response time to a hundredth of a second. This first procedure required an average time of 25 min.

Two lists of words were prepared for each subject from the subject's response times for use in the two subsequent tests. The sixteen emotional test words to which the subject had given the longest response times and the sixteen neutral words to which the shortest response times had been given, were extracted from the complete list of ninety words. Two lists were then prepared, each consisting of eight of the emotional words and of eight of the neutral words, arranged in a random order. Words were apportioned to each list so that the mean response time to the emotional words in one list was approximately equal to the mean response time to the emotional words in the second list and similarly for the neutral words.

(iii) *First tachistoscopic presentation*

This test took place at a period of approximately a fortnight after the initial word-association procedure. Ninety cards were prepared for use with the tachistoscope. On each card, white in colour, one of the ninety initial test words was stencilled in black Indian ink in letters of $\frac{3}{16}$ in.

The tachistoscope used consisted essentially of a frame into which the card was inserted. In front of the frame were two gas-filled bulbs, the light from which was reflected by the card and focused by means of an anastigmatic camera lens with a variable speed shutter on to a ground-glass screen 4 by 5 in. in dimension. In front of the screen, which was hooded, was stretched a length of thin white cord. A black mark at the middle of the cord indicated that the projected word would be seen a quarter of an inch above the mark when the observer was correctly seated.

The subject was seated before the tachistoscope and the following instructions were read by the experimenter:

This is an experiment in seeing or perception. I am going to flash a five-lettered word for a fraction of a second on to this screen. The words will look like this (a sample card is shown). I want you to tell me what the word is when it appears on the screen just above the black mark. Before each time I flash a word on to the screen, I'll say 'ready', and two seconds later it will appear. It is quite likely that you won't be able to recognize the word the first time or first few times, but I want you to tell me what it is as soon as you can make it out. I would like you to tell me when you see even only a part of the word. It's all right to guess. Is there anything which you are not sure about? All right! We will adjust your position and then we will do two practice words to show you how easy it is. After that we will go straight on to the others.

The seating of the subject was immediately adjusted so that a sample word projected with a high intensity of illumination and of a long duration could be seen clearly and at a

height of $\frac{1}{4}$ in. above the black mark. The average intensity of illumination of the adapting field was 1·72 apparent foot-candles and of the test field 1·75 apparent foot-candles. Two practice words (one emotional and one neutral) were then shown in accordance with the following procedure: the card was inserted into the frame by the experimenter, who was hidden by the tachistoscope, and the intensity of illumination of the bulbs (indicated by a light-meter inside the apparatus) was adjusted with the aid of a 'variac' resistance to 1 foot-candle. The word 'ready' was called and the test word was flashed on to the screen 2 sec. later. All words were to be projected for a constant duration of $\frac{1}{25}$ sec. The practice word was projected in increasing steps of 1 foot-candle of intensity of illumination until the subject gave two consecutive correct verbal reports. This procedure was repeated for the second sample word. Following the practice words, the sixteen test words were shown with the same procedure. These words, however, were projected at an initial intensity of illumination of 10 foot-candles below the lowest value at which the subject had correctly recognized the practice words.

The exposure time of $\frac{1}{25}$ sec. was used to ensure that projected words would be exposed for a period below the critical duration time in order that the Bunsen–Roscoe law would hold and that the product of intensity of illumination and of exposure time could thus be assumed to be constant (see Woodworth & Schlosberg, 1954). With reference to the fore-period of 2 sec. used in the experiment, it is appreciated that this value, as determined by Woodrow (1914), cannot now be accepted without question. On the basis of the writer's previous experience, however, and in the light of the observations of Teichner (1954), it was decided to retain such a fore-period in the experiment.

(iv) *Second tachistoscopic presentation*

A fortnight after the first presentation to each subject of her sixteen test words, the second list of sixteen test words was given. The experiment took place under exactly similar conditions. As before, the subject was seated in front of the tachistoscope. The previous adjustments and instructions were repeated, and with the exception of the use of new practice words, the experiment was continued in exact accordance with the above procedure.

(v) *Results*

To determine whether there was a significant difference between the reaction times to the forty-five emotional words and the reaction times to the forty-five neutral words, the Mann–Whitney test was applied to the results of each subject. Only subject 2 (see Table 2) showed a significant difference ($P < 0{\cdot}05$). However, it can be seen from the

Table 2. *Mean reaction time of each subject to each of the two classes of words*

	Subject											
	1	2	3	4	5	6	7	8	9	10	11	12
Emotional words	1·27	1·54	1·47	1·07	2·26	2·06	1·73	2·65	1·52	1·71	2·41	1·85
Neutral words	1·19	1·19	1·49	1·04	2·33	1·95	1·66	2·57	1·38	1·58	2·19	1·75

table that, with the exception of subjects 3 and 5, all the subjects showed a slightly longer mean reaction time to the emotional words. The Wilcoxon matched pairs test (Mosteller & Bush, 1954) was therefore used to ascertain whether the reaction

times of the group as a whole to each of the two classes of words were significantly different from each other. A significant difference was found ($P < 0{\cdot}01$; smaller sum of ranks $= 5{\cdot}5$).

From each of the two tachistoscopic tests, two mean recognition thresholds were obtained for each subject (see Table 3); the mean recognition threshold for the eight emotional words shown ($E1$ and $E2$) and the mean recognition threshold for the eight neutral words shown ($N1$ and $N2$). The differences $E1 - N1$ (in the case of the first test) and $E2 - N2$ (in the case of the second test), indicate the amount by which the mean recognition threshold for emotional words exceeds the mean recognition threshold value for neutral words. A positive value for this difference would indicate that the mean recognition threshold for emotional words was greater than the mean recognition threshold for neutral words. A negative value for this difference would indicate the reverse.

Table 3. *Individual results in the first and second tachistoscopic tests*

	First tachistoscopic test			Second tachistoscopic test		
	Mean recognition thresholds (ft.-candles)*			Mean recognition thresholds (ft.-candles)*		
Subject	Emotional words ($E1$)	Neutral words ($N1$)	Difference ($E1 - N1$)	Emotional words ($E2$)	Neutral words ($N2$)	Difference ($E2 - N2$)
1	18·63	16·00	2·63	23·63	21·50	2·13
2	20·63	20·00	0·63	16·13	14·13	2·00
3	15·00	18·13	−3·13	20·38	22·75	−2·37
4	13·50	14·75	−1·25	12·38	13·88	−1·50
5	18·63	18·00	0·63	12·63	12·00	0·63
6	21·88	21·56	0·32	17·38	17·50	−0·12
7	13·88	12·38	1·50	14·63	13·50	1·13
8	16·88	16·56	0·32	11·75	12·00	−0·25
9	13·38	14·00	−0·62	14·63	14·50	0·13
10	15·00	12·38	2·62	21·00	18·00	3·00
11	16·63	15·13	1·50	15·63	15·75	−0·12
12	16·13	14·50	1·63	17·00	15·63	1·37

* Recognition thresholds represent in each case the mean recognition threshold value of eight words.

The results of this experiment would allow the possible application of at least three forms of treatment. (1) For each test, it can be determined whether the mean values of columns $E1$ and $N1$ or $E2$ and $N2$ are significantly different from each other. (2) It can be similarly determined, for each subject, whether the mean recognition threshold for the eight emotional words is significantly different from the mean recognition threshold for the eight neutral words. (3) The agreement—in both extent and direction—between the performance of subjects in the first tachistoscopic test and their performance in the second test may be computed. Procedures (1) and (2) would allow application to either a single test or to the results of both tests. It was initially decided, on the basis of the observations in § II (vi), that individual performance in both tests could be most fruitfully examined by using the third treatment. The use of procedure (1) would consist of treating the results as a whole. The differing performance of individual subjects would obscure the opposite performance of other subjects and would leave an insignificant value, although certain subjects behaved consistently in one direction. Procedure (2) could not be used satisfactorily, as the initial selection of such common words—differing to such a slight extent in emotional connotation—would place an undue strain on the

possible contribution of factors of emotional selectivity as measured by this type of experiment. Treatments (1) and (2) were applied, however, to the results of both the single tests and to the results as a whole. No significant differences were found with the use of either non-parametric statistics or transformed data. The third treatment was therefore carried out. The correlation between the two difference columns $E1-N1$ and $E2-N2$ was computed by means of Kendall's rank correlation method (Kendall, 1948). This correlation was found to be highly significant ($\tau = +0{\cdot}73$, S.E. $= 0{\cdot}22$). The value obtained using Spearman's coefficient was $+0{\cdot}85$ ($P < 0{\cdot}001$, $t = 5{\cdot}07$, D.F. $= 10$).

Fig. 1. Differences between mean threshold values for emotional and neutral words of individual subjects in both tests.

Fig. 1 shows the performance of each subject in the first and second tachistoscopic tests. The length of a column indicates the difference between mean threshold values for emotional and neutral words and thus corresponds to the value of either $E1-N1$ or $E2-N2$, as shown in Table 3. A column above the horizontal axis indicates that this difference was positive. Each pair of columns represents the results of one subject in the first and second tests. The ordering of columns corresponds to the position of the subjects in Tables 2 and 3.

It can be seen from an inspection of Fig. 1 that the subjects showed the characteristic perceptual defence and sensitization in relation to emotional words referred to in our survey. Six of the subjects repeated their performance of possessing higher thresholds for emotional than for neutral words. Two of the subjects showed lower thresholds for emotional than for neutral words on the two occasions.

IV. Discussion

In this experiment, and in several others, the objective criterion of the emotional connotation of a word has been the duration of reaction time. Strictly speaking, the present experiment has only shown that higher thresholds (and the reverse) are obtained for

words of a long reaction time; that the words were actually emotional has not been unequivocally demonstrated. Our present knowledge indicates that two determinants are mainly responsible for the duration of reaction time: the associative strength of the stimulus word and its emotional effect on the subject. The more recent work on this question appears to indicate that with normal subjects it is the associative strength of the stimulus word which is the greater determinant of the length of reaction time.

It was shown earlier that the mean reaction times of the group as a whole to each of the two classes of words were significantly different. Our extraction, in the case of each subject, of the emotional words giving long reaction times and of neutral words of short reaction times may conceivably have resulted in words which differ mainly in their associative strength. We cannot assume that the use of a word count effectively controls individual differences in factors of word frequency. Laffal (1955), using words of roughly equivalent frequency from the Thorndike and Lorge count, has found a high correlation between long reaction times to the stimulus words and the number of different responses made to these words by the group as a whole. We should therefore expect, on the basis of Solomon & Postman's (1952) experiment, that words giving a longer reaction time would require higher recognition thresholds. Such an interpretation does not satisfactorily explain the performance of those subjects who show a greater sensitivity to stimulus words which earlier gave a longer reaction time.

The writer wishes it to be noted that all the studies which show the importance of the associative strength of words have been carried out with groups of subjects. That it may not be learning factors alone which determine recognition thresholds or length of reaction time is indicated by the fact that the correlations found do not show complete agreement. This is shown even when information measurement is applied to the data. Certain of the subjects, therefore, may not be showing such relationships. Further, the influence of emotion or the law of effect, whatever theory is held about their mode of action, must clearly have a place in determining the associative strength of words measured in frequency experiments.

How may we interpret, then, the findings of this and related experiments? The measurement of psychogalvanic responses, not only during the tachistoscopic session, but also during the word-association test, would be a valuable adjunct to an interpretation of the mechanisms involved. Again, the experiments which appear to show the persistence of individual forms of 'cognitive control' in varying situations, or which show a correspondence between clinical typologies and the differential perception of emotional words are suggestive. However, in this type of experiment, much requires to be understood regarding the processes which are responsible for these characteristic forms of perception.

Perhaps one of the most valuable ways of understanding our findings lies in the 'hypothesis' theory of Bruner (1951*a*) and his collaborator, Postman. Without entering into a detailed presentation of this particular theory, we can examine its application in a simplified form to a naïve laboratory subject. When the subject is seated before the tachistoscope and is informed that five-lettered words are to be shown, we can consider that she possesses a particular hypothesis or expectancy that common words will be presented. When, as in the preliminary stages of presentation, the letters are lacking in clarity and several different interpretations are possible, the strong hypothesis that common words are being shown is relatively easily confirmed. The thresholds for common words will therefore be lower. If, on the other hand, the subject holds the strong

hypothesis that the experimenter will actually show uncommon or taboo words, we can expect that these words will be more easily recognized.

This type of theory, with its various refinements, works well when the words are sharply distinguishable in character, as in the usual type of experiment. In the present experiment, however, the differences between emotional and neutral words are very small indeed. Bruner's theory is to be considered as a conceptual model, and thus can also be applied to our findings. When the term 'hypothesis' or 'expectancy' is used synonymously with the word 'set' and a definite physiological basis is presumed for these processes, as in Allport's (1955) interpretation, we find certain difficulties. In other words, can a theory which presumes a physiological basis for 'hypotheses' account for the different thresholds for words which differ to such a slight extent in nature?

Certain considerations are suggested by the results of the present study. There is a need to ascertain just what place factors of frequency have in determining perceptual thresholds when the material used is highly common. There is a need to find some method of studying such factors which is independent of the use of frequency counts. Again, one must agree with Vernon (1955) that one of the greatest defects of this form of research is its undue concentration on artificial situations. The use of concepts such as 'hypothesis' or 'set' has the greatest advantages for an understanding of the present type of experiment; not only because it gives a possible measurable basis to the processes involved, but because it suggests a means of combining the large amount of American work in this field with the peculiarly British contribution in the field of schemata.

The writer wishes to acknowledge the valuable assistance received from Mrs Pearl Segal, Miss Thelma Veness, Mr J. Brown and Mr N. Brown. He particularly wishes to express his gratitude to Mr B. M. Foss for his advice and criticism throughout the experiment and to Mr E. Wasservogel for his careful construction of the apparatus.

References

Allport, F. H. (1955). *Theories of Perception and the Concept of Structure.* New York: John Wiley; London: Chapman and Hall.

Bartlett, F. C. (1932). *Remembering.* Cambridge University Press.

Bitterman, M. E. & Kniffin, C. W. (1953). Manifest anxiety and perceptual defense. *J. Abnorm. (Soc.) Psychol.* **48**, 248–52.

Blake, R. R. & Ramsey, G. V. (eds.) (1951). *Perception: An Approach to Personality.* New York: Ronald Press.

Bricker, P. D. & Chapanis, A. (1953). Do incorrectly perceived tachistoscopic stimuli convey some information? *Psychol. Rev.* **60**, 181–8.

Brown, D. R. & Adams, J. (1954). Word frequency and the measurement of value areas. *J. Abnorm. (Soc.) Psychol.* **49**, 427–30.

Bruner, J. S. (1951*a*). Personality dynamics and the process of perceiving. In Blake, R. R. & Ramsey, G. V. (eds.). *Perception: An Approach to Personality.* New York: Ronald Press.

Bruner, J. S. (1951*b*). One kind of perception: A reply to Professor Luchins. *Psychol. Rev.* **58**, 306–12.

Bruner, J. S. & Postman, L. (1947). Emotional selectivity in perception and reaction. *J. Personality*, **16**, 69–77.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychol. Rev.* **62**, 193–217.

Delucia, J. J. & Stagner, R. (1954). Emotional vs. frequency factors in word-recognition. *J. Personality*, **22**, 299–309.

Dukes, W. F. (1955). Psychological studies of values. *Psychol. Bull.* **52**, 24–50.

Eriksen, C. W. (1952). Defense against ego-threat in memory and perception. *J. Abnorm. (Soc.) Psychol.* **47**, 230–5.

ERIKSEN, C. W. (1954). The case for perceptual defense. *Psychol. Rev.* **61**, 175–81.
FREEMAN, J. T. (1954). Set or perceptual defense? *J. Exp. Psychol.* **48**, 283–8.
HOWES, D. (1954). A statistical theory of the phenomenon of subception. *Psychol. Rev.* **61**, 98–110.
HOWES, D. H. & SOLOMON, R. L. (1950). A note on McGinnies' 'Emotionality and perceptual defense'. *Psychol. Rev.* **57**, 229–34.
HOWES, D. H. & SOLOMON, R. L. (1951). Visual duration threshold as a function of word probability. *J. Exp. Psychol.* **41**, 401–10.
HUMPHREY, G. (1951). *Thinking: an Introduction to its Experimental Psychology.* New York: John Wiley; London: Methuen and Co.
JAENSCH, E. R. (1930). *Eidetic Imagery.* London: Routledge and Kegan Paul.
KENDALL, M. G. (1948). *Rank Correlation Methods.* London: Charles Griffin.
KING-ELLISON, P. & JENKINS, J. J. (1954). The durational threshold of visual recognition as a function of word-frequency. *Amer. J. Psychol.* **67**, 700–3.
KLEIN, G. S. (1954). Need and regulation. In Jones, M.R. (ed.), *Nebraska Symposium on Motivation.* Lincoln: University of Nebraska Press.
KÜLPE, O. (1904). *Bericht über den 1st Kongress für experimentelle Psychologie*, pp. 56–68.
LACEY, O. W., LEWINGER, H. & ADAMSON, J. F. (1953). Foreknowledge as a factor affecting perceptual defense and alertness. *J. Exp. Psychol.* **45**, 169–74.
LAFFAL, J. (1955). Response faults in word association as a function of response entropy. *J. Abnorm. (Soc.) Psychol.* **50**, 265–9.
LAWRENCE, D. H. & COLES, G. R. (1954). Accuracy of recognition with alternatives before and after the stimulus. *J. Exp. Psychol.* **47**, 208–14.
LAZARUS, R. S. (1954). Is there a mechanism of perceptual defense? A reply to Postman, Bronson and Gropper. *J. Abnorm. (Soc.) Psychol.* **49**, 396–98.
LAZARUS, R. S. & LONGO, N. (1953). The consistency of psychological defenses against threat. *J. Abnorm. (Soc.) Psychol.* **48**, 495–9.
LAZARUS, R. S., ERIKSEN, C. W. & FONDA, C. P. (1951). Personality dynamics and auditory perceptual recognition. *J. Personality*, **19**, 471–82.
LAZARUS, R. S. & MCCLEARY, R. A. (1951). Autonomic discrimination without awareness: a study of subception. *Psychol. Rev.* **58**, 113–22.
LUCHINS, A. S. (1951). An evaluation of some current criticisms of Gestalt psychological work on perception. *Psychol. Rev.* **58**, 69–95.
MCGINNIES, E. (1949). Emotionality and perceptual defense. *Psychol. Rev.* **56**, 244–51.
MCGINNIES, E. (1950). Discussion of Howes' and Solomon's note on 'Emotionality and perceptual defense'. *Psychol. Rev.* **57**, 235–40.
MURDOCK, B. B. (1954). Perceptual defense and threshold measurements. *J. Personality*, **22**, 565–71.
MOSTELLER, F. & BUSH, R. R. (1954). Selected quantitative techniques. In Lindzey, G. (ed.), *Handbook of Social Psychology.* Cambridge, Mass.: Addison-Wesley.
OLDFIELD, R. C. (1954). Memory mechanisms and the theory of schemata. *Brit. J. Psychol.* **45**, 14–23.
POSTMAN, L. J. (1953). The experimental analysis of motivational factors in perception. In Brown, J. S. *et al., Current Theory and Research in Motivation: A symposium.* Lincoln: University of Nebraska Press.
POSTMAN, L., BRONSON, W. C. & GROPPER, G. L. (1953). Is there a mechanism of perceptual defense? *J. Abnorm. (Soc.) Psychol.* **48**, 215–24.
POSTMAN, L., BRUNER, J. S. & WALK, R. D. (1951). The perception of error. *Brit. J. Psychol.* **42**, 1–10.
SOLOMON, R. L. & HOWES, D. H. (1951). Word frequency, personal values and visual duration thresholds. *Psychol. Rev.* **58**, 256–70.
SOLOMON, R. L. & POSTMAN, L. (1952). Frequency of usage as a determinant of recognition thresholds for words. *J. Exp. Psychol.* **43**, 195–201.
TEICHNER, W. H. (1954). Recent studies of simple reaction time. *Psychol. Bull.* **51**, 128–49.
THORNDIKE, E. L. & LORGE, I. (1944). *The Teacher's Word Book of* 30,000 *Words.* New York: Teachers College, Columbia University.
VERNON, M. D. (1955). The functions of schemata in perceiving. *Psychol. Rev.* **62**, 180–92.
WOODROW, H. (1914). The measurement of attention. *Psychol. Monogr.* **17**, no. 5 (whole no. 76).
WOODWORTH, R. S. & SCHLOSBERG, H. (1954). *Experimental Psychology.* London: Methuen and Co.

(*Manuscript received* 6 *July* 1955)

PUBLICATIONS RECENTLY RECEIVED

The Child's Conception of Space. By Jean Piaget and Barbel Inhelder. Translated from the French by F. J. Langdon and J. L. Lunzer. London: Routledge and Kegan Paul. 1956. Pp. xii + 490. 42*s*.

It is greatly to the credit of British psychologists and publishers that so many of the major works of Jean Piaget have first appeared in translation in this country, rather than the U.S.A. Mr Langdon and Mr Lunzer must be congratulated on the excellence of their translation of the often difficult and abstruse argument of this book. It is to be hoped that due recognition may now be given to the outstanding contribution made by Piaget and his collaborators to our understanding of the nature of the thought processes as they develop from infancy through childhood to the adult stage.

The Child's Construction of Reality described the investigation of the infant's immediate perception and his early understanding of the nature of objects. *The Child's Conception of Space* (first published in 1948 as *La Représentation de l'Espace chez l'Enfant*) proceeds to demonstrate how this stage is succeeded at a considerably later age, first by an intuitive understanding of the manner in which objects are related to one another in space. This is followed by a pre-logical grasp of spatial relations within projective space to the viewpoint of the observer; and finally by completely logical concepts of integration within the co-ordinates of Euclidean space. The child works out the successive stages for himself by means of appropriate motor activity, the images of which are integrated within ideational schemata, leading finally to what Piaget calls 'operations'.

Though they may find the treatment by Piaget and Inhelder of this development in terms of mathematical logic difficult to comprehend, all psychologists will welcome the ingenious experimental demonstration of the manner in which the various spatial concepts evolve. Thus it is shown that the earliest notion of spatial relations is the topological, since this constitutes the simplest form of spatial organization. Experiments on tactile perception of shape and on drawing simple and complex forms demonstrate that the understanding of proximity and separation precedes that of simple geometrical shape, especially where angles and inclined lines are involved. Experiments on threading rows of beads show the grasp of order; and on the tying of knots, that of surrounding or enclosure. The concept of continuity, which draws together these notions, appears in the synthesis of points to make a continuous line.

The development of ideas about projective space is then demonstrated in the correct judgement of parallel lines and of the shadows and cross-sections of geometrical solids. But, most interesting of all, is the ability to envisage the appearance of landscape scenes from different points of view. For it is only when the child has grasped the effect of difference in viewpoint that he begins to understand how the disposition of objects in space may be related to the position of his own body; and hence the notion of projective space in general. But the final stage, the understanding of Euclidean space, comes only when the measured distances of objects from each other and from the observer can be related within the framework of a universal system of co-ordinates. This is demonstrated by two most important experiments. The first, on estimating the position of the surface of a liquid in a tilted vessel, shows that the child does not reach an immediate understanding of the horizontal and vertical co-ordinates of space from his direct perception of the horizon and of vertical objects, but develops it gradually up to about the age of 9 years. The second, on the reproduction in a scale plan of the positions of scattered objects, demonstrates that the total co-ordination of spatial positions within a single whole is not attained till the age of 10–11 years.

Piaget's experimental techniques have been criticized for their lack of experimental control. In fact, the descriptions of the behaviour of children and their answers to questions in these experimental situations do not prove that all children proceed through these stages of development; nor can exact age norms be derived from them. Further experiment with larger numbers of children is needed; or, even better, a longitudinal study in which the same children are followed up at successive ages. But these observations, even as far as they go, add enormously to our understanding of the conceptual development of children; and of the extreme complexity of spatial concepts and ideas which to the adult seem so simple and obvious. It is impossible in the space of a review to demonstrate satisfactorily the subtlety and profundity of Piaget's argument. Thus it is to be hoped that all psychologists will study it for themselves.

M. D. VERNON

Risk and Gambling: the Study of Subjective Probability. By JOHN COHEN and MARK HANSEL. London: Longmans, Green. 1956. Pp. x+153. 14*s.*

In recent years guessing behaviour has become a phenomenon of interest to psychologists. In part this interest has been aroused by a concern for psychophysical methodology, and in part also by an increased awareness of the problems involved in the study of human thought processes. This book is a very timely addition to the literature available. It describes a series of experiments of a guessing or gambling nature, carried out by the authors mainly on children. Thus it has additional value in improving our knowledge of children's concepts of causation. The experiments are extremely ingenious, and will undoubtedly be repeated as class demonstrations in many university courses and other lectures in general psychology. Psychologists and statisticians who are exercised about problems of response dependencies in those situations demanding repeated observations from the same individual will find their fears confirmed, and be driven further to elucidate 'psychological randomness' and to devise procedures to minimize its unwanted effects.

It seems churlish to point out defects in the book, but, as it is the first publication on these problems to reach the non-specialist public, some comments must be made. Excellent though the experiments are, there are many others, mainly in the American literature, which could well have been cited in support. A few of the earlier post-war references are given in footnotes. Further, there is no reference to the nineteenth-century statistical literature on the relation between probability and belief (cf. especially Venn) which is so relevant. Perhaps most surprisingly of all in work which is primarily an exercise in the psychological significance of Bayes's Postulate, Bayes is never mentioned. But with those defects (which are to some extent offset by due appreciation of Piaget's approach), this work remains a valuable and important contribution to one of the central problems in psychology.

J. W. WHITFIELD

The Normal Child and some of his Abnormalities. By C. W. VALENTINE. London: Penguin. 1956. Pp. 291. 3*s.* 6*d.*

Though the subtitle of his book is 'A General Introduction to the Psychology of Childhood', Prof. Valentine has set himself the very specific task of considering what Prof. Mace in his foreword describes as 'a question of great concern to parents and teachers to-day', namely, 'Is this child "normal" or shall I take him to the clinic?' In the course of offering a basis on which this question may be answered, Prof. Valentine considers children of all ages from infancy to early adolescence and in a wide variety of activities and human relationships. In spite of this, it is doubtful whether his central question can provide enough integration wholly to justify the subtitle. The book certainly cannot be called 'popular' in any pejorative sense; as one would expect, it is written with scrupulous concern for the citing of evidence and the weighing of arguments. But, as the author says himself, it 'only seeks to be descriptive and explanatory'. It may reasonably be argued that a general introduction should be more largely concerned with general theory than this book, given its first and guiding aim, is able to be.

However, as a source of reassurance to the anxious parent, troubled by scraps of psychological information—or misinformation—and by vague notions about 'neurosis', its value should be considerable. It offers for such a parent the opportunity to develop perspective and to acquire standards of comparison or, to use the word suggested by the title, 'norms'. One of the difficulties in any attempt to do this is, of course, this very word 'normal', with its varying and conflicting uses and connotations. Prof. Valentine begins his book with a direct attack on the question of what may be meant by the 'normal child'. 'The word "norm"', he says, 'means a rule, standard or average.' The trouble, however, is that the first and second of these words share all the ambiguities of 'norm' itself. The 'standard' may be the average or it may be the ideal to be aimed at, in the sense in which standards are set. This distinction between the statistical and the evaluative 'norm' is a common source of confusion, since an approach to the standard in one sense is generally a departure from the standard in the other. And a corresponding but perhaps even greater difficulty is caused by the word 'abnormal', for the negative introduces the problem of deciding at what point of remoteness from the statistical or from the evaluative 'norm' behaviour is to be considered statistically or evaluatively abnormal. It might have been helpful if Prof. Valentine had considered these issues more fully than he does. He may well have felt that any very detailed discussion would be out of place in a book of this kind, but, in its absence, there must be a certain risk of recurring perplexity for any reader who does not come to the work with a prior awareness of the main distinction.

MARGARET C. DONALDSON

Psychology: the Fundamentals of Human Adjustment. (Third edition.) By NORMAN L. MUNN. Boston: Houghton Mifflin Company. 1956. Pp. xvi+542. $5.75.

Dr Munn has reorganized his popular text-book and brought it thoroughly up to date. Unexpectedly this has led to a reduction in size effected by widespread pruning and by giving up the former arrangement of chapters in divisions each with its own introduction. The twenty-four chapters of the second edition have been reduced to sixteen rather longer ones without any noticeable contraction of coverage. In some directions there has been an extension to cover formerly neglected topics, such as fatigue.

A new sequence of topics has been adopted in order to shift neurology to the end of the book and to '...capitalize, at the outset, upon the interest of students in individual differences...'. The new chapters in turn deal with *The Scope and Methods of Psychology*, *Individual and Group Differences*, *Intelligence*, *The Motivation of Behavior*, *Emotional Motivation*, *Frustration and Conflict*, *Personality*, *The Learning Process*, *Foundations of Learning*, *Remembering and Forgetting*, *Thinking*, *Attending and Perceiving*, *Vision*, *Hearing and the Other Senses*, *Neural Foundations of Behavior*, and finally, *Working Efficiently*. Difficulties raised by this sequence are most obvious in the discussion of emotion where the Bard–Cannon theory precedes any mention of neurones and conditioning appears prematurely in connexion with John Watson's theory. Most of the chapters are only loosely integrated. The order in which they are set down need not dictate the order of a course based on the book.

Since more than one-sixth of the references and more than half of the suggestions for further reading date from 1950 or later, this edition is well up-to-date. It is to be regretted that psychological writings should become obsolete as quickly as they do.

The 35-page glossary is a useful new feature but would be improved by pruning non-technical words, such as *rage*, *reproof*, *research*, *rivalry*, *dreaming* and *love*. The last is defined non-technically as 'A feeling of attachment or affection...not necessarily sexual', but the only indexed reference is to the word as used by John Watson. Some of the definitions of important terms are weak. Emotion is defined as 'A condition underlying such experiences and actions as occur in fear, rage and the other so-called emotions'.

Generally the text is carefully written, but now and again loose expressions occur. We are told on p. 243 that statistical analysis of some experimentally observed differences shows that 'the chances are almost 100 in 100 that they result from different motivation, and not from chance factors'. A picture of a distorting room bears the caption 'How our eyes can deceive us'. Giving examples of constancy Munn writes '...the square table top is continually seen to be the same square...' in spite of changes in the retinal image.

The book sets out to attract students. Illustrations are so plentiful that the chances of opening it and being confronted with two pages of unbroken type are about 3 to 1 against. The judicious banishment of calculational procedures in statistics to an appendix has further reduced the chance of a student's encountering aversive stimulation. Most of the illustrations are well chosen, but a few, e.g. a comic strip of Poincaré going through the four stages of creative thinking, are only distracting and may embarrass the serious student.

The text preserves the same flavour as earlier editions despite some shift in emphasis from rats to people and from the physiological to the social aspects of the basic processes. Experimental psychology and the S.–.R. approach are most fully represented, but Dr Munn is able to take up a wide variety of points of view with sympathy. At times there is an irritating nebulousness about the discussions of basic processes, but this is an almost inevitable result of eclecticism. Admirers of the earlier editions will not be disappointed with the changes that have been made. In many ways the book remains the best of its type.

Some degree of colour shock may be induced by the new jacket. R. L. REID

A Primer of Freudian Psychology. By CALVIN S. HALL. London: Allen and Unwin. 13*s.* 6*d.* (British edition; first United States edition published in 1954.) 1956. Pp. xii+137.

The first thing that must be said about this book is that it would be difficult to improve on its workmanship—I might almost say 'craftsmanship'. Prof. Hall has set himself the task of giving an account in Freudian terms of the development of normal personality, as opposed to abnormalities. Most attempts of

this kind are either superficial, short and implausible or thorough, technical and long; but this one manages to be short, plausible and yet technically accurate. The secret of this achievement seems to lie partly in the systematic order that Prof. Hall has worked out for introducing the various interlocking concepts without repetition or tiresome cross-references. But it also lies in his knack of graphic yet economical explanation. Here is the Freudian theory of a frequent cause of neurotic anxiety, epitomized in eighteen words: 'The id says "I want it"; the superego says "How horrible"; and the ego says "I am afraid".' Could anything be shorter or clearer?

Nevertheless, the author is very careful to include only what can be found in Freud's published works, and to exclude early views which he can be shown to have abandoned. This is so conscientiously done that I was proud to be able to find a few doubtful passages, which I dare not do more than question. Did Freud really believe that 'There would be no psychological development if every time the baby began to feel the tension of hunger it was immediately fed and if all of the other excitations arising in the body were similarly discharged by the co-operative efforts of parental care and inborn reflexes'? Is Freud responsible for the over-simplified physiological account of asthma on p. 89? Is 'predicate-thinking' really a Freudian term? And does not the 'primary process' include condensation and displacement as well as the production of 'a memory-image of an object that is needed to reduce a tension' as Freud implies, for example, on p. 52 of his *Outline of Psychoanalysis* (1940)?

These are minor details, however. I have a more serious doubt about the conception of this book. Some of the Freudian theories which it outlines have been considerably modified by Freud's own followers, such as Klein and Fromm (unless one defines 'follower' as 'one who makes no modifications'). It would have been useful if Prof. Hall had outlined these changes as clearly as he has dealt with the original. The alternative was to recognize that an exposition of Freud and nothing but Freud is of historical interest, and to treat it historically by tracing the development of his theories (a task which Dr Ernest Jones has not rendered unnecessary). But—probably for the sake of simplicity—Prof. Hall has tried to make his book a *cento* of Freud's last thoughts, and has thus reduced its usefulness. The book is otherwise so good that this is doubly unfortunate, and I hope that the second edition will contain at least footnotes to indicate the main neo-Freudian developments.

NIGEL WALKER

Freud and Religious Belief. By H. L. Philp. London: Rockliff. 1956. Pp. xi + 140. 18*s*.

Not unlike Shelley, Swinburne, Nietzsche and Charles Bradlaugh, Freud, despite his genius for interpreting the activities of the unconscious mind, failed to realize that a man's atheism can be as emotionally motivated as a theist's religious beliefs. The existence of God cannot be proved, neither can we find sufficient conclusive arguments to prove the opposite. But the arguments of an atheist should be sympathetically examined and this is what Dr Philp proceeds to do in his book. Having dealt with what appears to be the emotional basis of Freud's unrepentant atheism he carefully scrutinizes the evidence and conclusions of the great analyst as stated in *Totem and Taboo*, *The Future of an Illusion* and *Moses and Monotheism*. Freud's attack on religion has long been thought to be as illogical as it is emotional, and because of its prejudicial origin it is inevitably distorted. Dr Philp submits that the distortion arose out of Freud's reaction against anti-Semitism in the first place and was exaggerated by his limited knowledge of anthropology, by his mishandling of data that was second-hand, and by his eagerness to force all religious beliefs into the categories of illusion, thereby committing the fallacy of generalizing from a few particulars. 'He attempted to demolish religion by tracing the motives on which he thought it was based in particular cases and his own atheism was open to precisely the same form of attack'. Freud has given invaluable service in helping us to discriminate between false and valid experiences, even in religion, which the author frankly acknowledges. 'The unconscious desires and wishes give rise to the process of projection and rationalization... in forming his religious beliefs man is largely under the influence of these motives and processes'. But neither Freud nor anyone else has been able to prove the non-existence of that which is claimed to be experienced whether it be an elf or God. Where he principally failed was in narrowing his interest 'until it is concentrated on what appears to fit in with psychoanalytical theory'. Dr Philp's purpose is to make clear Freud's determination to get rid of Christianity because of its anti-Semitic association, and that the reasonableness of his attack could only be sustained by the overthrow of all religion on the ground of wish-fulfilment. Nevertheless, it is not logically sound to try and rebut the position of the atheist by presenting or asserting the claims of the theist, and Dr Philp comes perilously near to committing this fallacy in pp. 80–6. He assumes the existence of God to prove that projection can substantiate the claim that 'God is primarily seeking man rather than man creating the image of

God'. However, he has done an excellent work in revealing the flimsy data on which Freud has based his unbelief. 'I believe that Freud's case against religion in general was unsound and reflected his own opinions rather than the findings of psychoanalysis'. Clearly he shows that anyone wishing to proclaim atheism would be well advised not to quote Freud.

W. E. SARGENT

Introduction to Psychopathology. By LAWRENCE I. O'KELLY and FREDERICK A. MUCKLER. London: Staples. 1956. Pp. xi+704. 50*s.* net.

This book is the second edition of one first published in America in 1948. It has been written jointly by the Professor of Psychology at Illinois University, and his Research Assistant. It is designed for advanced university students, and each chapter concludes with a list of topics for discussion and suggested reading and bibliographical sources. The book also contains a useful glossary and a long bibliography. This collection of references is, as might be expected, excellently complete and the book is useful for this alone. Some of the text, however, is rather elementary, as when the reader is warned not to confuse factor analysis with psycho-analysis.

The book is divided into four parts dealing respectively with: introduction to basic concepts, the problems of disordered behaviour, the causes of disordered behaviour, and theories and treatment of disordered behaviour.

Despite its title this book could equally have been called a text-book of abnormal psychology. It is disappointing that so little space has been given to treatment and prophylaxis and none whatever to techniques of investigation.

Nevertheless, this book will undoubtedly be useful to students, although the professional clinical psychologist still waits for a comprehensive text-book on techniques of investigation, methods of treatment and their underlying rationales. However, it is probably too early to expect this.

RALPH HETHERINGTON

Prenatal and paranatal factors in the development of childhood behavior disorders. M. E. ROGERS, A. M. LILIENFELD, and B. PASAMANICK. Baltimore: Johns Hopkins University. No date given. Pp. 157. No price given.

This monograph describes an ambitious and carefully controlled investigation into the frequency with which abnormalities of pregnancy and delivery occurred in the histories of 1151 white and non-white elementary schoolchildren in Baltimore, who were considered by their teachers to suffer from behaviour disorders.

When all the children with behaviour disorders are grouped together, a greater frequency of abnormalities in the prenatal and paranatal periods can be demonstrated among them than among a series of 902 controls. A high proportion of these disorders occurred in the children who had been born prematurely. These findings are of great interest, but from the practical point of view the study would have been of more value if the authors had attempted to ascertain which particular forms of behaviour disorder were associated with which abnormalities in the prenatal and paranatal periods. Unfortunately, the only group they were able to study in this way was the one they labelled 'Hyperactive, confused-disorganized cases' in which most workers already accept the aetiological importance of birth injury—an opinion confirmed by the writers.

As it stands, the investigation merely shows that there is an apparent association between the occurrence of behaviour disorder and a previous history of abnormal prenatal or paranatal periods in a large series of cases. No directly causal effect can be attributed to the abnormalities of the prenatal or paranatal periods on the evidence offered, and it would require much closer study and more detailed classification of the forms of deviant behaviour observed before any symptoms or group of symptoms could be considered to be due to the abnormalities of the prenatal or paranatal periods which they noted. Such a study would obviously be well worth while, but it is doubtful if it could be done without the investigators having much closer clinical contact with, and knowledge of, their patients than is evidenced in this exhaustive statistical analysis.

T. T. S. INGRAM

Suicide in London. By PETER SAINSBURY. 1955. London: Chapman, Hall. Pp. 116. 15*s*.

This ecological study of suicide develops the sociological approach pioneered by Durkheim and by workers in the U.S.A., including Cavan, Faris and Mowrer. It is a major contribution, providing a valuable body of facts and a model of scientific method, and should be read by all those interested in social psychiatric investigations.

Designed to test a carefully formulated hypothesis, two complementary investigations were undertaken: (1) a statistical correlation of the suicide rates of the London boroughs with selected indices of their social characteristics, using largely the Registrar General's publications and Smith's 1932 *New Survey of London Life and Labour*; and (2) an analysis of the records of 409 suicides reported to the North London coroner in 1936–8.

Sainsbury has shown clearly that the suicide rates of the London boroughs differ from one another, and that these differences correlate with certain social characteristics of the districts—their social isolation, mobility (whether of class, occupation or neighbourhood), and disorganization (as manifest in divorce and illegitimacy rates). These correlations, and the detailed findings of this and similar surveys, are fully and critically discussed.

No attempt is made to claim too much for the findings of the second half of the investigation. Sainsbury rightly points out that the social and statistical approach must be complemented by studies of individual cases, and that the best of coroners' reports provide only very partial information about the personal factors involved. He stresses the need for more intensive psychiatric and psychological studies of attempted suicides. A useful bibliography lists the relevant literature only to 1950.

I. R. C. BATCHELOR

Man, Motives and Money: Psychological Frontiers of Economics. By ALBERT LAUTERBACH. New York: Cornell University Press; London: Cumberlege. 1954. Pp. xiv + 366. $5.

Prof. Lauterbach, a teacher of economics, takes his stand on the belief that people in their economic roles function no differently than in their other roles; in particular, they are always complex figures whose attitudes, motives and behaviour are affected by their culture, society, personality and all the rest. This is not to admit, he goes on, that psychology can replace economics; but isolation and analysis of 'economic' behaviour can only be useful if it does not ignore the psychology of personality, of perception, of unconscious processes and of communication. Nor is it possible for economics to go its own way and then 'apply' the finished findings of psychology (as one has heard economists argue); it is the approach to the exploration of economic processes that must be psychological.

The first topic taken up is business behaviour in our society, and the factors that influence it and its underlying motivation. The range of discussion is wide, including motives of competition, price determination, industrial location and investment decisions, myths and symbolism in business life, and emotional factors in beliefs about free enterprise. Summing up, what evidence there is, is to the effect that important influences of a non-economic type 'are not an exception but a rule in determining the ways in which business is started, expanded and currently done'. The point is that these influences are cultural, social, psychological, almost anything but what might be called economic. Thus in price determination it seems that a price 'often has little relation to the costs of manufacture; sometimes it is vaguely connected with whatever is considered a fair price'—an intriguing link with the fair day's work of the Hawthorne bank memory!

The second main topic is the relation of economic instability to personal insecurities and especially the implications of prosperity without stability. It is argued that socio-economic instability affects personal insecurities so that these react again on socio-economic attitudes and behaviour. The third part of the book strikes out even more widely (one almost said wildly) into the psychological implications of economic reform. The reason why people have rarely followed prescriptions for economic improvement, the writer argues, is that these prescriptions seldom took account of the attitudinal and emotional implications. The danger is, however, that reform programmes may be put forward which, though purporting to be economically based, are simply exploiting emotional needs in the service, for instance, of political ideologies. It seems to him an important task to help people distinguish between genuine economic requirements and

'pseudo-economic rationalizations of unfulfilled emotional needs'. A major difficulty, it seems to the reviewer, is to know where one rationalization ends and another begins.

The author's basic point of view is not new. As long as fifty years ago economists were asking for a closer relation with psychology. On the other hand, those economists who are interested in the traditional theoretical formulations will at once admit that they deal with almost pure abstractions; they will add, however, that they like it that way. The book itself is uneven, and obviously it must be rather speculative, although the author is the first to stress that evidence is not available to support more than a very few of the generalizations. He paints economic psychology with a rather broader brush than does, say, Katona. But his aim is to offer challenging hypotheses and stimulating research suggestions. His belief is that only from multi-disciplinary research can their verification come. And when he argues that other social sciences (e.g., cultural anthropology) have something to say which may prevent new attacks of ethnocentrism in economic thought one cannot, as a psychologist, control a rather guilty start.

R. M. McKENZIE

How people vote. By MARK BENNEY, A. P. GRAY and R. H. PEAR. London: Routledge and Kegan Paul. 1956. Pp. xii+227. 25*s*.

This is the final report of a study by a group of London School of Economics sociologists of electoral behaviour in the Greenwich constituency during the 1950 General Election. They used the Lazarsfeld technique of repeatedly interviewing a representative panel of electors, and analysed the way the votes went in terms of such variables as age, class and interest in politics. They discuss the stereotypes of the political parties and the attitudes of members and others to national issues.

All this is a great improvement on the usual agglomeration of electoral gossip (plus the cube law) which normally appears in this field—although it must be said that about one-third of the book is taken up with just such detail. But it is in the remaining two-thirds that the 'meat' of the work lies.

From the social psychologist's viewpoint, there are two main criticisms of the material presented. The results are given in a large number of direct comparisons of percentages. (61% of the middle class, but only 44% of the working class, described themselves as being interested in politics; so did 47% of those aged 65 and over, as against 49% of those aged 21–39.) One would have liked a rather more sophisticated (and less tedious) way of dealing with the data. Even in the present format there should have been a definite statement as to the statistical significance of each difference discussed. This is so even though we would, of course, agree with the authors that it is sometimes profitable to discuss statistically non-significant differences. The second point is more important and concerns the rather slipshod way in which some of the opinion data are treated. For example, the authors seem quite unaware of the techniques for constructing an attitude scale, and the use of their 'Index of Political Opinion' is quite illegitimate. The results they obtain with it (e.g. that two-thirds of Conservatives are strongly Conservative whereas only a third of Labourites are strongly Labour) may be quite misleading, since no account is taken of the bias of their measure.

This is a difficult book to evaluate. Judged in terms of the literature of what has come to be called 'psephology', it is undoubtedly the best work to have appeared in this country so far. And here the matter might rest were it not for the high standards that the book itself sets up in a brilliant first chapter (by Benney) on the purpose and perspective of this type of study. We were promised a comparison between 'the public universe of politics and the millions of private universes of the electors', or, put in other terms, a working through from political science to social psychology. By these standards the book is disappointing. The reason for the failure lies, I think, in the fact that to attain an interdisciplinary end, a project must itself be interdisciplinary. This sort of study calls out on almost every page for the work of, say, psychologists and social anthropologists; it must be a co-operative venture if it is not to be shrugged off by the non-participating disciplines as superficial, pedestrian or misguided. At this level, this review is merely a psychologist's shrug.

DEREK PUGH

Interpretation of Sociometric Choice Status. By ÅKE BJERSTEDT. Lund: Gleerup. 1956. Pp. 408. 40 Sw. Cr.

In recent years, sociometric techniques have been used increasingly to explore social phenomena in a great variety of situations. The now common procedure of requiring the individual to show his willingness to interact with certain individuals as well as his unwillingness to interact with certain others, has, in its

simplicity, a methodological attractiveness. But all too frequently, it would seem, the surface simplicity of the method has masked its complex character. In the absence of established methods and rigorous interpretative examination of results, many sociometric studies have yielded obscure and often conflicting conclusions. What would seem necessary if sociometry is to come of age is to attempt to gain a unanimity in what sociometry studies, a scientific rigour in the methodology, and a sound theoretical framework for what is purposed by using the techniques. In the book reviewed here, a serious attempt is made to attain certain of the objectives.

The author discusses and reports tests of a series of common interpretative assumptions concerning sociometric choice status. In his task, he sets out first to discover the definition and delimitation of sociometry. His questionnaire to different authors who have written on the subject reveals much terminological confusion but also several ways in which agreement might be reached. In the first part of the book it is also necessary for him to discuss the general principles for collection and analysis of socio-preferential data and to decide on their suitability in testing the assumptions.

Parts Two and Three of the book are devoted to the testing and reporting of outcomes of the assumptions that sociometric status scores are: (*a*) relatively unaffected by preference methods; (*b*) relatively unaffected by choice aspect; (*c*) relatively constant over a period of time; (*d*) accompanied by the individual's awareness of his own relative sociometric situation; (*e*) correlated with measures of general 'adjustment' and 'efficiency' of the individual; and (*f*) amenable to similar interpretation in different countries. In Part Two an internal analysis of socio-preferential variables is made from a broad sample of Swedish schoolchildren. Attention focuses in Part Three on a smaller sample of subjects representing extreme socio-preferential positions. An extensive comparison is made of the socio-preferential variables with a series of extra-socio-preferential variables. The outcomes of the tests reveal a fair amount of non-specificity with respect to methods, choice aspects and test moments. There is evidence presented of a certain awareness by the individual of choice status received. Several clear extra-socio-preferential correlations are found with respect to specific aspects of 'adjustment' and 'efficiency'. Finally, in those areas where comparison is possible, the Swedish results show fundamental similarities with Anglo-Saxon reports.

Throughout the discussion and the two studies, there is ample evidence of a vigorous scientific inquiry. The work of most sociometric researchers is given a careful examination and the results are compared with the author's work. A continual attempt is made to effect a synthesis or to improve on the techniques. Such an approach points out many areas that require further—and often immediate—attention. Those using sociometric techniques will find many valuable suggestions in the book. For those who have not used sociometric methods and are attracted towards their use, this book should be required reading, for, as few other more general works on sociometry have done, there is illustrated here the essential 'frontier' character of sociometric research, the still unfinished state of the techniques, the dangers of over-simplification and naïve interpretation, and the tentative character of most results. The author's concluding remark is that 'there is still much to be done'. This would seem true not only for the methodology of sociometry but for all that is purposed by using the techniques.

A. R. MACKINNON

Soviet Attitudes toward Authority. By MARGARET MEAD. London: Tavistock. 1955. Pp. 148. 21*s*.

The subtitle of this book, 'An Interdisciplinary Approach to Problems of Soviet Character', indicates its aim, which is to trace the impact of Bolshevik ideas and policies on the older Russian character and its consequent transformation. The account of traditional Russian character structure is largely based on Gorer's exposition of the consequences of swaddling, although this is not directly mentioned. The continued persistence of some of the older traits is shown to result in contradictions and conflicts, whence arises the necessity of constant rigid control and the stress on self-criticism.

Most of the work is devoted to a discussion of the way in which these fundamental attitudes influence leadership roles and the relationship between the leaders at various levels of the hierarchy and the masses. There is much emphasis on the now odious personality cult, a fact which is understandable, as the book first appeared in America in 1951.

The chief, and inevitable, weakness of the study lies in its reliance on indirect materials, mainly pre-1950 Soviet publications and interviews with emigrants whose reliability is hard to assess. Given these limitations, the book is stimulating and certainly offers valuable insights into past official Soviet attitudes. A crucial question is implicit in a statement by the author herself: 'These materials provide no direct

information on the extent to which those to whom they are addressed believe what they read or on the extent to which they act upon it.' It is possible that much of the material used represents a paper curtain which screens rather than reveals the actual attitudes and behaviour of Soviet citizens. Reports of reactions to the Khrushchev speech at the XXth Congress tend to point in this direction. However, until it becomes practicable to investigate attitudes directly, judgement will have to be suspended.

G. JAHODA

Papers of the Michigan Academy of Science, Arts, and Letters. Vol. XLI. Michigan: University Press; London: Cumberlege. 1956. Pp. viii+368. 80*s*.

Only the last four papers of this volume are of direct interest to psychologists. One paper concerns a longitudinal study of factors likely to predict children's readiness for school. Another treats style of emotional expression as related to social class, suggesting that lower-class expression is more immediate and bodily and less indirect and symbolic than middle-class, and that this is reflected in the forms of psychopathology to which members of different social strata are prone. The third reports techniques of rating the ambiguity or structuredness of therapeutic situations. The last investigates cognitive reorganization following electroconvulsive treatment and concludes that, during recovery, orientation returns first to personal identity, then to place and other people, and last to time. The papers' brevity precludes their giving more than a hint of these four research programmes being conducted at Michigan, and it would be extravagant of any library to buy the volume only for their sake. It is to be hoped that fuller reports of these programmes will soon be available in a form more suited to the interests and purses of psychologist readers.

I. M. L. HUNTER

Horse Psychology. By MOYRA WILLIAMS. London: Methuen. 1956. Pp. xvi+194. 16*s*.

This book comes of letting a clinical psychologist loose in the paddock. The expected result might be a psychology of the centaur, but Mrs Williams succeeds in animalizing her beloved subjects with commendable detachment. Her explanations of horse behaviour are eclectic, leaning rather heavily on animal ethology. The horse has been used in so few experiments that most of the evidence quoted comes from other animals; however, some ingenious pilot experiments on her own horses are described and wittily illustrated. The book is ideal for adult education classes, and many psychologists will read it with pleasure and profit. Perhaps Mrs Williams might reach an even wider audience if an erratum slip were inserted: for 'horse' read 'child'. As it is, it will surely be placed in many Christmas and birthday lists.

BRIAN FOSS

Medizinische Psychologie. (11th edition revised.) By E. KRETSCHMER. Stuttgart: Thieme. 1956. Pp. viii+382. DM. 29.60.

Most text-books of psychiatry bear the author's stamp to a degree quite unknown in science or in the better established medical specialties. Kretschmer's views have always been particularly strongly held, and this book is very much an expression of his own views and a presentation of his own system. No account is taken of criticisms, no mention is made of empirical work disagreeing with Kretschmer's own views. In the psychiatric field, these views are usually interesting, stimulating and derived from close clinical experience. As suggestions for research, they will be of interest to the psychologist. As a text-book, however, the work is too one-sided to be recommended.

To the psychologist, the last few chapters dealing with personality tests, projective tests and psychotherapy, are probably the most interesting. Unfortunately, they are well below the standard of the rest of the book. Kretschmer is clearly covering a vast territory without any specialized knowledge of the literature or even of the principles according to which validity and reliability are to be judged. It is perhaps characteristic that hardly any references are given and that those which are given are either out of date or refer to Kretschmer's own work. Hardly any of them are in English.

Kretschmer, in his introduction, considers his treatment of psychological tests to be rigorous, logically structured and orientated towards essentials. It is difficult to find any evidence for this belief in the chapters themselves.

H. J. EYSENCK

Das Gemüt. By S. STRASSER. Utrecht/Antwerpen: Het spectrum; Freiburg: Herder. 1956. Pp. xix+291. No price given.

Though the philosophical analysis of emotion, feeling, affect and related phenomena has received considerable attention, there appears to be no systematic phenomenological survey to offset the increasing weight of the 'positive sciences' in this field. It is the author's aim to present such a doctrine, using phenomenological evidence as a basis for a holistic-structural and non-causal interpretation.

A wide area of emotional and feeling tone is covered. Although relying extensively on traditional, particularly scholastic and modern (Hartmann, Scheler) philosophy, it is heartening to see to what extent outspoken psychological theories (Krüger, MacDougall, Sander, Stern) are preferred to existentialistic ones (Merleaü-Ponty, Sartre). That is, unfortunately, where this certainly stimulating and painstaking treatise rests. Instead of making the most needed attempt towards reconciliation between philosophical psychology and empirical evidence, the latter is too easily dismissed, using only a very insufficient and unsystematic reference to experimental and theoretical literature.

J. C. BRENGELMANN

Der Aufzähl-Test (*The Naming Test*). By ADOLF BUSEMANN. 1955. München/Basel: Reinhardt. Pp. 174. No price given.

The subtitle is 'Investigations into the World of Experience of Healthy and Neurotic Children'. The world here means 'things': the realm of object relationships. This, in the author's opinion, throws light on the child's relationship to the world in the wider sense, including personal relationships.

The instruction requires the naming of 'visible' objects for 10 min. Abstract and invisible things are not scored, nor are repetitions of the same word. Scoring is based on the total number of words, but attention is given to a number of features such as the distribution of the words within the 10 min.; the 'profile' of the graph produced; length of the time elapsing between instruction and the first word, and between words or groups of words; number of repetitions of the same word; groups and classifications according to subject or complex; the number of things present and not present, etc.

Results show that records are fairly constant for the same individual, but are sensitive to changes in mental state and to organic changes, e.g. brain injuries.

As a frame of reference for interpretation, the author gives a detailed and comprehensive description of personality assessment, remarkable for a well-balanced attitude to the different schools of thought, and for honesty about the possibilities and limitations of the test.

This test seems a very promising diagnostic instrument, well worth developing. It seems to yield valuable information at the least expense in time and effort. It touches on the deeper layers of the personality and lends itself to interpretation in terms of dynamic psychology, without disturbing the subject.

The numbers on which the present investigations are based are too low to draw conclusions as to the validity of the test. It is to be hoped that the author will develop the pattern of scoring, so as to provide an instrument for the clinical psychologist.

MARTHA VIDOR

Genie, Irrsinn und Ruhm. By WILHELM LANGE-EICHBAUM. Basel: Reinhardt. 1956. Pp. 628. Swiss Fr. 36.70.

This is the fourth edition of a classic which was originally published in 1927. In the second and third editions, the main text had been left unchanged and a large number of additions, footnotes and postscripts were added. The author had hoped himself to rewrite the work completely, but was prevented by his recent death from doing so. W. Kurth undertook the task and has carried it out in a manner which would appear to have been in line with the original author's intentions. Some sections have been eliminated altogether, the list of references has been brought up to date and considerably enlarged, and all the previous addenda have been incorporated in the text. The result is a very much bigger book, which, none-the-less, has retained its readability and its general organization. It has also retained its stress on the sociological conception of genius.

It would not be reasonable in a brief review like this to summarize or criticize a classic which ought to be familiar to all psychologists. If it were not for the language barrier, the interest of the topic and the

high quality of the style would together have ensured a much wider knowledge of the main thesis of this book. It is to be hoped that this new edition will be made available to English-speaking readers, although the fact that hardly any of Lange-Eichbaum's references are to the English literature may speak against this.

H. J. EYSENCK

OTHER PUBLICATIONS RECEIVED

Acculturation: Critical Abstracts, North America. By B. J. SIEGEL (Ed.). Stanford: University Press; London: Cumberlege. 1956. Pp. 231. 32*s.*

Therapeutische Psychologie ihr Weg durch die Psychoanalyse. By W. M. KRANEFELDT. Berlin: de Gruyter. 1956. Pp. 152. DM. 2.40.

A Handbook of Medical Hypnosis. By G. AMBROSE and G. NEWBOLD. London: Baillière, Tindall and Cox. Pp. 256. 21*s.*

The Political Role of Women. By M. DUVERGER. Paris: UNESCO. 1955. Pp. 221. 13*s.* 6*d.*

Emile for Today: the Émile *of Jean-Jacques Rousseau, Selected, Translated, and Interpreted.* By W. BOYD. London: Heinemann. 1956. Pp. 198. 10*s.* 6*d.*

Karl Marx: Selected Writings in Sociology and Social Philosophy. By T. B. BOTTOMORE and M. RUBEL. London: Watts. 1956. Pp. 268. 21*s.*

Political Theory. By G. C. FIELD. London: Methuen. 1956. Pp. 297. 18*s.*

The Primary Psychiatric Syndromes. By D. L. MOODY. Bristol: Wright. 1956. Pp. 356. 37*s.* 6*d.*

Educational Psychology in the Classroom. By H. C. LINDGREN. London: Chapman and Hall (Wiley). 1956. Pp. 521. 40*s.*

The Practice of Psychiatry in General Hospitals. By A. E. BENNETT, E. A. HARGROVE and B. ENGLE. California: University Press; Cambridge University Press. 1956. Pp. 178. 30*s.*

A Psychiatric Approach to Institutional Work with the Aged. By H. TURNER (Ed.). New York: Community Service Soc. No date. Pp. 44. $00.60. (Mimeographed.)

Prospero and Caliban: a Study of the Psychology of Colonisation. By O. MANNONI. Translated from the French by P. POWESLAND. London: Methuen. 1956. Pp. 218. 22*s.* 6*d.*

The Education of Young Children. By D. E. M. GARDNER. London: Methuen. 1956. Pp. 118. 7*s.* 6*d.*